

Lie detectors tell too many lies

The popularity of lie-detector tests in the United States is distressing and dangerous. Their use by the federal government has set a bad example to other employers.

WHEN it seemed in the nineteenth century that there was nothing that enough steam, electricity and gears could not do, and when even Mary Shelley wrote books like *Frankenstein* to prove it, it may have seemed natural that a machine to which enough wires were attached could tell if a man was honest. We may no longer believe that a room full of high-voltage transformers can bring the dead to life, but the idea of the lie detector, a machine that can probe a man's inner heart, seems just too alluring. In the United States, there is now even a prime-time television show called "Lie Detector" which each week gives some convicted felon the chance to prove, by submitting to a polygraph test while millions watch, that he is an innocent man wronged. It is great entertainment. Unfortunately, thousands of high government officials, grocery store managers and police sergeants now see the lie detector as something more. Last year, they ordered more than one million polygraph tests in the belief that they would find the leaker of classified information, hire the honest employee or catch the thief.

In fact, all that they accomplished was a systematic invasion of privacy with a procedure only slightly more reliable than flipping a coin or examining a sheep's entrails. And most alarming is the federal government's uncritical rush to expand mandatory lie detector testing for federal employees — and to impose penalties upon those who dare to stand up for their rights and refuse to have their fate determined by a game of chance masquerading as science.

Professional polygraphers often claim that their results have been shown to be as much as 99 per cent accurate. But this turns out to be the result of uncritical and self-deluding experiments or, worse, outright manipulation of statistics. If one throws out tests run under artificial conditions and requires a minimum of double-blind methodology, the lie detector's much-vaunted accuracy rapidly diminishes to little better than chance. One of the best recent studies used polygraph charts taken from subjects interrogated for actual crimes. Fifty subjects subsequently confessed to the crime in question; they were matched with 50 subjects who had been interviewed for the same crimes and, thus, exonerated by the confessions. All, of course, denied committing the crime under the polygraph test. The charts and records of the polygraph tests were submitted, on a blind basis, to six polygraph experts for scoring. They were able to tell who was the liar and who was the honest man only 69 per cent of the time. Similar controlled studies have confirmed this result.

Endless games can be played with statistics. Some of the highest accuracy rates reported often turn out to be only percentages of guilty subjects scored "deceptive" on the polygraph test. Of course, one does not even need a polygraph machine to make that figure up to 100 per cent — just say everyone is guilty. The "98 per cent" score often cited becomes much less impressive when matched with the false positive rate — the number of innocent subjects scored "deceptive" — in the same test: 55 per cent. The average of 98 per cent and 45 per cent is back to 71 per cent.

The inherent inaccuracy of the polygraph test, even under the best of conditions, means that many innocent people are needlessly brought under suspicion. Even worse, the tests are frequently used under far from ideal conditions. All the careful studies of polygraph accuracy involved tests in which answers to specific questions ("Did you steal \$50 from the desk drawer?")

were sought. Last spring, President Reagan ordered all federal employees to submit to polygraph tests in the course of investigations of leaks of classified information. These tests tend to be "screening" tests — a sort of dragnet that even the Federal Bureau of Investigation does not allow. The Department of Defense, until Congress ordered it to stop, was planning a massive expansion of its use of the polygraph in an even broader and less technically defensible context: screening employees and job applicants on such matters as drug use, homosexual activities, connections with communist groups and the like. The National Security Agency and the Central Intelligence Agency are said to make extensive use of such testing of employees.

Employers outside the federal government have now joined the bandwagon. Retail stores and banks in particular seem to like the idea. The questions asked of employees include: "Do you intend to stay with this employer?", "Are you satisfied with your job?", "Have you ever intentionally underpriced merchandise?". In this setting, the polygraph is not even being used as a detector of deception: it is being used as a device to intimidate the subject into potentially embarrassing or damaging admissions. Stressful responses during the test (for that is really all the machine measures) are used as the jumping-off point for detailed grilling of the subject. No studies have ever suggested that the polygraph has any scientific validity in such screening tests.

But even when used in the verified mode of operation, the polygraph test subjects innocent people to embarrassing ordeals. The state-of-the-art method is known as the "control question technique"; relevant questions ("Did you steal the \$50?") are matched with control questions ("Have you ever stolen from an employer?") designed to provoke a greater response from the innocent subject. The control question is crafted to get a rise out of the innocent person; incidentally, in answering it, he may be forced to reveal some damaging information.

The greatest danger of a reliance on polygraph tests, apart from the invasion of privacy it entails, is the false sense of security it affords. David Lykken, a University of Minnesota psychologist who has studied — and criticized — polygraph testing, says it is not difficult for a trained person (or for that matter a psychopath) to beat the machine. Lykken tells of a prisoner with whom he corresponded — a convicted murderer — who, after reading Lykken's critiques, trained 27 fellow inmates who were accused of breaking prison rules (and who told Lykken's correspondent that they had indeed done so) to beat the polygraph test that the prison used to adjudicate such matters. By such stratagems as biting their tongues when the control question was asked, 23 of them passed. The prisoner who wrote to Lykken, incidentally, had been convicted on the basis of a polygraph test — prosecutors said they would drop the charges if he passed the test, provided that he allowed the test to be used as evidence against him if he flunked. He flunked. Two and a half years later, another man confessed to the murder, and he was released.

In 19 states and the District of Columbia, it is illegal to ask an employee to take a polygraph test. Legislation may soon be introduced in Congress to forbid polygraph testing of federal employees (the congressional prohibition on expanded Department of Defense use expires on 15 April 1984). It is much needed as a step in the right direction of stopping the perpetuation of foolish injustice. □



UK civil science

Mason report proposes yet more reform

THE British preoccupation with the administration of publicly supported civil research continues. Sir Ronald Mason, appointed in August to carry out a one-man inquiry into the working of the relationship between British government departments and the titularly autonomous research councils (see *Nature* 11 August, p.476), this week delivered the firm opinion that the departments are inept at telling what research they need. Many of those who opposed the Rothschild reorganization of research in 1971, which among other things established that government departments should behave as "customers", commissioning research from "contractors" represented by the research councils, will now be saying "We told you so".

The Mason inquiry was commissioned by the Advisory Board for the Research Councils (ABRC), whose chief function is annually to recommend to the Department of Education and Science the partition of the budget available for civil science between the five research councils. Mason has ensured a sympathetic reception from his sponsor by urging that the advisory board should have more authority and responsibility. His overall objective is to make the Rothschild principle function more effectively, not to abandon it.

Mason's chief complaint about the working of the Rothschild system is that it appears to have made little difference in the pattern of departmentally supported research except to diminish support for "strategic" research. But departments which are not the mainstay of single research councils (as is the agriculture ministry for the Agricultural Research Council (ARC), for example), have since 1977 taken advantage of the Rothschild prescription to reduce their total support for research council work by about a quarter, from £26.6 million to £19.9 million (in 1982-83 prices).

The report of the inquiry also points out that all government departments have declined to follow the Rothschild prescription that something like 10 per cent of the cost of research commissioned from research councils should be regarded as a surcharge to support cognate strategic research — the agriculture departments on the basis that funds for the purchase of commissioned research with ARC has never exceeded 55 per cent of the council's total budget, compared with the 70 per cent transfer originally recommended.

What has gone wrong? The Mason report says that there is no reason to doubt the validity of the Rothschild recipe, but that the Rothschild proposals have not

been fully implemented. In particular, government departments have not functioned in relation to the research councils as "informed partners" — but nevertheless there has been an increase in the bureaucracy and in administrative costs of the kind that led, three years ago, to the decision that the Department of Health should hand back its "Rothschild money" to the Medical Research Council (MRC).

Not all the fault lies in Whitehall, however. Mason points out that the research councils are less flexible than they should be in accepting research contracts because they employ most scientists on terms based on those common in the scientific civil service, in which job security until retirement age is the general rule. Mason recommends that the research councils might usefully "examine" this practice, saying that its abandonment would give the research councils more flexibility and a stronger flow of new entrants to their research laboratories.

Apparently in passing, but pointedly, the Mason report says that both ARC and the Natural Environment Research Council (NERC) should worry about the large proportion of their research carried out in their own laboratories, and says that at many of these institutions there is a need for "reorientation, amalgamation, privatization or transfer to an appropriate government department".

Mason's recipe for reform, that ABRC itself should be given more influence, derives from his suggestion that govern-

ment departments should not waste further effort on building up an infrastructure for providing intelligent advice on research programmes but that, instead, they should recruit part-time scientific advisers from among people of "stature". Then it would be possible for ABRC to function as a means of holding the ring between conflicting research needs, to which end it should have "a more full-time chairman" and an adequate secretariat. One of the functions of such a strengthened board would be to sanctify agreements between departments and research councils on substantial research programmes, whereupon the partner council would be allowed to get on with the work under supervision from the board.

Mason also raises the possibility of closer liaison between ABRC and the British Government's other chief source of advice on civil science, the Advisory Council on Applied Research and Development. More cross-membership is a must, even a merger of the two committees should not be ruled out. With that suggestion, Mason (whose report is nowhere over-respectful of his terms of reference) has come as close as he might have thought it prudent to questioning the government's firm decision a year ago that there should not be a government department concerned with research and development in both science and technology.

Although formally a report to ABRC, most of Mason's recommendations are in reality addressed to government departments and even the cabinet. Almost inevitably, however, his comments on NERC and ARC will influence the outcome of the ABRC recommendations on the division of the annual science budget, due to be published before the end of the year.

John Maddox

Ocean research out?

As a result of government pressure for increased efficiency, the UK Institute of Oceanographic Sciences (IOS) faces the possible closure of its sites at Taunton, Devon, and Bidston, near Liverpool, and the merger of the two institutes in new buildings near the present head office at Wormley in Surrey. This is one of the more severe options being costed by the Natural Environment Research Council (NERC), which runs the institute. The proposals will be assessed by the council at its meeting early in December.

The annual expenditure of IOS is about £2.3 million net. In 1982-83, about £4.3 million worth of research was commissioned by government departments. According to the director, Dr A.S. Laughton, the level of government commissions has held up quite well (unlike that in other areas supported by NERC) but the time-scales over which forward

planning is possible within such programmes has decreased. As a result of NERC's drive for staff reduction, IOS is losing more than 3 per cent of its staff each year through natural wastage.

Of the three institutes within IOS, Wormley is by far the largest, with a staff of about 250 and with research programmes spanning physical, chemical and biological oceanography and geology. Taunton and Bidston (with about 50 and 70 staff respectively) work in more applied areas such as tidal computations (Bidston) and applied wave research and sedimentation (Taunton).

None of the options now being considered by NERC includes deliberate cessation of these activities, but there are many potential difficulties. For example, the tidal computation group at Bidston employs part-time members of staff who are both hard to replace and unlikely to want to move to the more expensive "gin and tonic" belt of Surrey.

Philip Campbell

Spacelab

Shuttle ready for launch again

Washington

THE maiden flight of Europe's Spacelab, cancelled last month because of last-minute doubts about the safety of the shuttle's solid rocket boosters, is to go ahead on 28 November, the National Aeronautics and Space Administration (NASA) announced last week. The European Space Agency (ESA) agreed to the new date only after receiving assurances that four European experiments that will lose a significant amount of data because of the postponement will be able to fly again by mid-1985. They are:

●A metric camera experiment designed to test the mapping capability of high-resolution photography from space. A large-film mapping camera is to be mounted at the optical window in the Spacelab module to discover whether photographs taken in orbit could provide a practical alternative to conventional mapping techniques. Because of the delay, cloud and snow cover over Europe and the low angle of the Sun, five out of 40 targets are likely to be missed.

●An infrared grill spectrometer designed to examine the atmosphere of the Earth at altitudes between 9 and 93 miles will lose about a third of its data because of orbital mechanical problems and because of the changed lighting conditions in November. NASA has promised that the instrument will fly on a later Spacelab mission in mid-1985.

●An experiment designed to take 2,000 photographs of OH (oxygen-hydrogen) emissions in the atmosphere at an altitude of 53 miles will lose at least a third of its data because of the absence of orbital night-time. The instrument, which could make important contributions to research on cloud structures in the high atmosphere, will be reflown in mid-1985.

●An experiment using a very wide field camera to make a general ultraviolet survey of the Milky Way and the region of the Sun will lose about 10 per cent of its data because of the extra daylight. The experiment will be reflown on the third Spacelab mission in November 1984. One of NASA's own experiments, far-ultraviolet astronomy using the FAUST telescope, will face similar problems and is to fly again in mid-1985.

Two spectacular space plasma physics experiments planned by NASA will also be seriously affected by the absence of dark, and are to fly again in mid-1985. One would use particle accelerators to fire high-intensity electron and ion beams into space. Particle beam injections create small-scale artificial auroras and trace magnetospheric processes. The other uses a low-light-level television and photometer to produce images of faint atmospheric emissions.

Peter David

European information technology

Research brief delayed to Athens

ESPRIT, the £850 million European strategic research programme for the information industries, appeared last Saturday to be tottering into the abyss of the European budget problem — the question of who pays what subsidy to Europe's over-productive farmers. The trouble is that both the farming subsidy and half the cost of Esprit (the rest would come from industry) are paid out of the same European Community coffers.

At last Saturday's research ministers' council in Brussels, Britain and West Germany (the only net contributors to the EEC budget) declined to support Esprit as an extra item. The other eight members agreed to a slightly trimmed budget (£800 million over five years), although France had earlier suggested halving the sum, but the deadlock over British and German support means that Esprit will be referred to the European Summit meeting at Athens on 6-7 December.

At Athens, Esprit will no doubt be used by Britain and West Germany as a bargaining card. Both countries already have substantial national research programmes in information technology, and so have less to gain from Esprit than do other countries, with the possible exception of France. Nevertheless, feeling is widespread in Europe that Esprit is a good programme, defined in detail by industry rather than from a Brussels office. (Although Inmos, the British designers of the "transputer" computer-on-a-chip, have struck a jarring note by describing Esprit as "too research oriented". The company also says it would be dangerous to be involved if Esprit were more applied, for fear of losing secrets.)

But Inmos apart, given Esprit's generally wide support, there would be immense chagrin if it fell at its last political hurdle. Commissioner Etienne Davignon, the architect of Esprit, has said that if the European summit fails to agree, and Esprit is lost as a separate programme, the money will be found within the existing Brussels research budget of £290 million a year. Since Esprit would cost the Commission £85 million a year, this solution would mean a reduction in the existing research and development budget (mostly energy research) of nearly 30 per cent. Other Brussels officials, however, are taking a more sanguine view. "I think the prime ministers want to take the credit of launching Esprit themselves" said one, who expected the summit to agree the programme. If Esprit goes ahead, the Commission will be ready to launch into the first year's work. This year, a £6.5-million pilot programme of 36 projects has been set in motion, and a detailed list of projects for the main programme will be ready on 1 December, in time for the summit.

The procedure already used for selecting pilot projects will be used in subsequent

years of Esprit, the research ministers decided last Saturday. The procedure takes only a matter of a few months from the outline proposals to the selection by the Commission of the best project. This is race-track speed for an inter-governmental body, but so far there have been few difficulties; in particular, there has been little acrimony about selection, despite obvious competition among European companies, a Commission official said on Monday. When programmes get down to details "it is pretty obvious who should do what, say in very large scale integrated circuits or software". Another potential point of friction is between Esprit and national programmes (such as Britain's Alvey directorate); but that will be ameliorated by a government overseeing body, the spokesman said.

Robert Walgate

Italian research

More sometimes means less

Milan

RESEARCH and development spending in Italy has risen above 1 per cent of gross national product (GNP) for the first time this year (to 1.3 per cent) Professor Ernesto Quagliariello, the president of the Italian research council (CNR), has announced in his annual report on Italian research. Total spending this year of 7 million million lire (£2,930 million) represents a 23.5 per cent rise on 1982, in current lire, to be set against an inflation rate of only 7 per cent.

Much of the increase is in applied research, with public and private companies contributing 52 per cent of the total spent, as against only 48 per cent last year. Basic research (by CNR and by the atomic energy authority ENEA) enjoyed an 18.6 per cent increase to a total of 3.3 million million lire (£1,380 million), but this is still not enough to guarantee a sufficient level of productivity, Quagliariello says.

Moreover Italy's massive government-funded applied research programme of 3.2 million million lire (£1,340 million), agreed under "law 46" in February 1982, has now run out of cash. During its 2-year existence, "law 46" spent 500,000 million lire (£210 million) in research contracts between industry and university, 1.2 million million lire (£500 million) in applied research in nationalized industry, and 1.5 million million lire (£630 million) on innovation support. Good money while it lasted, but now, according to research minister Luigi Granelli, a re-funding of applied research is absolutely necessary. Thus, a battle between applied and pure seems likely in the definition of the 1984 budget, due next month.

Paola de Paoli

Grenadian medical school

Offshore refuge for US students

Washington

THE 650 students at St George's University School of Medicine — whose presence on Grenada was used, at least for a few days, as justification for the United States invasion of the tiny Caribbean island — are members of a rather mediocre group of perhaps as many as 15,000 US citizens who are studying medicine abroad after having been rejected by US schools.

During the 1970s, more than a dozen such medical schools were set up in Mexico, the Caribbean and the Philippines to cater for such people. The schools are not accredited, have no entrance requirements other than ability to pay the tuition fees, and are attended overwhelmingly by North Americans.

St George's differs from other such schools only in the relative success of its students in returning to medical practice in the United States. Of the 218 St George's students who last year took the test set by the Education Commission for Foreign Medical Graduates (ECFMG) — a requirement for foreign-trained Americans wishing to obtain residencies at US hospitals — 79 per cent passed. The average success rate on the test is only 30 per cent.

But like five other "offshore" medical schools investigated by Congress's General Accounting Office in 1980, St George's

was criticized for having inadequate clinical training, poor laboratory facilities and too few qualified teachers. And, true to its type, St George's was founded in 1976 by an American entrepreneur — Charles Modica, himself a rejected medical school applicant. Modica, who is 36 years old, continues to serve as "chancellor" of the school.

Many of the schools were founded at a time when applications to US medical schools were rising sharply. For some 50 years until the 1970s, applications had held steady at a rate of two for each available place. During the 1970s, that rose to three for each place. A congressional amendment requiring US medical schools that received federal support to admit a percentage of US students who had completed two years at offshore schools

also increased the popularity of the schools for rejected US applicants. The Caribbean and Mexican schools also benefited from a decision by Italy — once a popular choice for rejected US students — to require a proficiency in Italian.

Since then, however, Congress has repealed the "Guadalajara" amendment and applications to US medical schools have fallen to the previous 2 to 1 level. The popularity of the offshore schools may wane further with the forthcoming revision of the ECFMG test due next spring. The test, which has been criticized for setting a standard below that of the national board examinations that US medical students must pass, will be toughened and will add, for the first time, a requirement that students should demonstrate their practical proficiency.

The very need for the offshore medical schools has increasingly been called into question by recent government projections of a surplus of physicians in the United States.

Stephen Budiansky

Women in science

No home from home in Israel

Rehovot

SEX discrimination against academic women has now become an issue in Israel, where, as elsewhere, women academics are a rare breed. Here academic women tend to work in the natural sciences and medicine — at the Hebrew University of Jerusalem, for example, there are nineteen women in the natural sciences and medicine but only four in the social sciences and humanities. Dr Galia Golan, a fierce feminist and the first professor of political science at the Hebrew University, has a cynical explanation: in the natural sciences there are objective criteria for measuring achievement but in the humanities "subjectivity is a major factor", while promotion committees are predominantly male. Professor Golan, who is seeking to increase the number of women in senior academic positions, claims "four children and four books" to her credit, but has throughout her career forgone sabbatical leave abroad because her husband is a physician in private practice.

Sabbatical visits overseas, more necessary in the natural sciences than in other fields, are a constant difficulty. One woman scientist at the Weizmann Institute says that her recent two years at a British university was a "hardship for the whole family". She took her two daughters with her, but her husband joined them only for brief visits. Yet, she says, scientists who "are serious" about their work have no choice but to go abroad.

The Weizmann women are an extraordinarily homogeneous group according to Dr Hadassah Horn of the Hebrew University. Almost all earned their PhD degrees at about the age of 30, and have spent five years reaching the highest non-tenured rank of senior researcher. With

one exception, all have children.

On past form, further advancement to a tenured post will be difficult. Women are a quarter of senior researchers at the institute but only 7 per cent of associate professors and 4 per cent of full professors. Much the same is true at the Hebrew University, where women account for no less than 65 per cent of those in non-tenure track positions but for only 21 per cent of lecturers, 13 per cent of senior lecturers, 12 per cent of associate professors and 6 per cent of full professors.

In Professor Golan's opinion, the explanation for this imbalance is that the criteria for academic advancement, set by men, do not fully take account of women's need to miss time in their late 20s and early 30s, when they have young children to look after. Professor Golan is also annoyed by lesser issues. It has finally been agreed that female academics should be allowed to stay at work until they are 65, but she says there is still discrimination against women in conditions of employment. She wants women to be given priority when the universities are once again able to recruit staff, and says there will be no danger that women would take priority over more highly qualified men because "there are plenty of women just waiting for the opportunity".

Not all women academics agree. Professor Ester Samuel-Cahn of the Hebrew University thinks the shortage of women in senior academic posts is a consequence of women's deliberate choice of family responsibilities. Her colleague Professor Alice Shalvi, mother of six children, agrees that academic women must make a choice between "neglecting their children and working".

Nechemia Meyers

Soviet aid in Grenada

THE Soviet Government has been planning a major technical aid programme intended to make Grenada self-sufficient in food, *Pravda* reported last week. A feature article by a hydraulic engineer, S. Sergeev, who visited Grenada last winter, gave a glowing report of the intentions of the Soviet Ministry of Land Improvement and Water Management to help the Grenadians to design canals and reservoirs, sink artesian wells, construct irrigation works and plan the lay-out of fields.

The soil could thus be made to yield several harvests a year, Sergeev explained, and the Grenadians would then be able to "supply themselves with foodstuffs which are at present imported from other countries".

Sergeev did not explain how this assistance would be repaid. Soviet aid programmes to Third World countries do not normally take the form of free gifts which would "pauperize" the recipients — except in cases of disaster relief. The usual pattern, when Soviet know-how is used to develop a country's mineral or agricultural resources, is for the beneficiary country to repay the investment over a protracted period with a proportion of the increased production made possible by that aid.

Vera Rich

Radiological protection

UK inquiry into leukaemia cluster

AN inquiry into allegations that radioactive discharges into the Irish Sea might have caused excess cancers in the coastal region of Cumbria was announced in the British Parliament last week by Mr Patrick Jenkin, Secretary of State for the Environment.

The allegations were made in a Yorkshire Television documentary (see *Nature* 3 November, p.5) which claimed to have found evidence of increased incidence of leukaemias and some other cancers near the coast of Cumbria, in the vicinity of a fuel reprocessing plant at Sellafield (formerly Windscale) run by British Nuclear Fuels Ltd. The programme also drew attention to radionuclides from the plant effluents that were found in household dust.

Polish indictment

THE act of indictment against the four Polish scholars — Jacek Kuron, Adam Michnik, Henryk Wujec and Zbigniew Romaszewski —, who acted as “intellectual advisers” to Solidarity (see *Nature* 6 October, p.466) is causing considerable puzzlement in Warsaw legal circles. One eminent lawyer reportedly observed that it is unclear whether the authorities wish to sentence them to a term of imprisonment or put up a monument to them.

Among other activities, carried out within the framework of the Social Self-Defence Committee “KOR” (the forerunner of Solidarity), the four are charged with “various activities and public campaigns contrary to the existing legal order” in pursuance of the aims of “a struggle against reprisals undertaken on political, ideological denominational and racial grounds, and bringing help to people persecuted on these grounds; exposing violations of legality and helping the victims thereof; and providing support and defence for any social initiative designed to further respect for human and civic rights”.

The long-term aim of the group is defined as “parliamentary democracy”, by which, the indictment says, was meant a system based on a “widespread movement of ‘self-governments’ independent of the Party and the State authorities”, to be extended in the first instance to state employees, agricultural producers, students, scientists and those employed in cultural work, and territorial units. This aim, the indictment said, was tantamount to attempting to overthrow the system.

In support of these aims, the group is accused of circulating unofficial newspapers, organizing steering committees for a future independent trade union movement, setting up “Students’ Solidarity Committees” in the university and sponsoring the unofficial “Flying University”.

Vera Rich

The government’s review of the evidence is to be carried out by Sir Douglas Black, a former president of the Royal College of Physicians, who will be working in collaboration with health officers in the region. The fact of a local increased cancer incidence appears not to be in dispute: it is the interpretation of that fact that has caused controversy. The village of Seascale, named in the programme as having an increased incidence of childhood leukaemia, has rapidly expanded its population in recent years, and the Yorkshire team has not divulged the details of calculations purporting to show that the incidence is very unexpectedly high.

The difficulty of interpretation is illustrated by the observation that cancer incidence over a large area of Cumbria in the 1970s was significantly less than expected. Dr Peter Tiplady, community health adviser to the West Cumbrian Health Authority, is working on a breakdown of cancer statistics by towns and villages, and only then will it be possible to see whether Seascale is an isolated phenomenon. Cancer “clusters” have only recently become a focus of interest for epidemiologists. The theory necessary to estimate their significance is being investigated by the National Radiation Protection Board (NRPB) in conjunction with epidemiologists at the University of Oxford.

BNFL has stressed that dosage to individuals is thought to be ten to one hundred times less than would be necessary to

account for the Seascale leukaemia figures. The programme suggested that household dust might constitute an exposure route that has not been considered, but NRPB is convinced that ingestion and inhalation of household dust are insignificant dose routes. It has found airborne radioactivity in the region to be unimportant from the viewpoint of radiological protection.

BNFL has complained to the Independent Broadcasting Authority about the research techniques used by the Yorkshire Television team (which allegedly included improper approaches to individuals) and will also be lodging a complaint about “one-sided advance publicity”. Highlights of the programme were given prominence in most newspapers in Britain before it was broadcast, and BNFL is annoyed that no other possible cause of the alleged cluster was mentioned.

One allegation in the programme that has particularly infuriated BNFL is the suggestion that radiation doses from its plant have been increasing. In fact, the company says, doses have decreased substantially in recent years and will be reduced further when an ion exchange effluent treatment plant is completed next year. Dosage to critical groups exceeded target recommendations for prolonged exposure of the International Committee on Radiological Protection (although not its dose limits) in 1981 and before, and the authorizing government departments have been “encouraging” BNFL to reduce its discharge levels. Draft authorizations now being considered will reduce discharges to one thirtieth of the present authorized level. The cost will be about £80 million.

Tim Beardsley

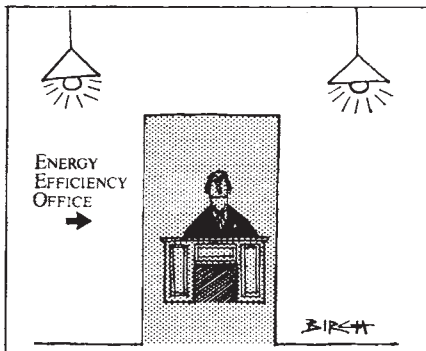
Energy conservation

UK sees light

THE British Government’s promised Energy Efficiency Office was launched with a flourish last week. Displaying the sort of evangelical zeal that cabinet ministers normally reserve for election times, Mr Peter Walker, Secretary of State for Energy, announced a “national crusade” to save money by promoting the more efficient use of energy. He aims to save one-fifth of the £100 million that Britain spends daily on energy, and promises that “the campaign will not end until we have succeeded”.

The Energy Efficiency Office is headed by Mr Bill McIntyre, who has worked in the oil industry and has previous departmental experience in conservation programmes. In its first year of operation, the office will have a budget of some £15 million, a 35 per cent increase on last year’s conservation expenditure. The publicity campaign will cost £3 million, less than the “Save it” campaign that was promoted in the 1970s, but officials hope that Mr Walker’s personal commitment will make up for that. Ministers will be visiting more than 40

towns with exhibitions advising consumers on how to reduce waste, and chairmen of electricity, gas and coal boards have pledged their support. The campaign is aimed at both industry and the domestic consumer.



The government’s declared strategy on energy conservation is that “realistic pricing” should be the main spur to efficiency, although even Mr Walker’s predecessor, Mr Nigel Lawson, recognized that there is a place for government encouragement. There is a common feeling that government incentives have recently been haphazard and largely ineffective. Now, Mr Walker, a notorious cabinet “wet”, seems keen to make amends.

Tim Beardsley

French universities

Laying down the laws

A NEW law for the French universities and *grandes écoles*, the first since the "democratic" reforms that followed the troubles of 1968, is working its difficult way through the French Senate. Since the Senate is profoundly conservative, and the architect of the new law — minister of national education Alain Savary — is "profoundly liberal" (according to one of his advisors), the passage is proving stormy.

Moreover, all might have gone better if M. Laurent Schwartz, renowned professor of mathematics at the Ecole Polytechnique in Paris, had not seen fit to publish a few weeks ago a book with the provocative title *To Save the Universities* (by implication, to save them from Savary). This has armed senators with useful facts — but, say its detractors, it is misleading about the true nature of Savary's law.

The problem may lie more in what the law leaves out, and leaves undefined, than in what it puts in. Take the "selection" of students for example. Lecturers are often appalled at the size of unselected first-year classes; Savary would increase them, regarding the first year as "orientation". He abhors "selection". But because of the realities of space limitations at the universities another clause in the law allows selection to take place in the second year, as it does now. To find the room to avoid selection it would be necessary to create new para-university institutes, whose students are "selected" or "oriented" towards them; these are recommended by M. Schwartz, but their foundation is also implied by M. Savary and his law. The two sides, in this as in many cases, use different rhetoric but arrive at the same conclusions.

Another point at issue is the question of "decrees", and their relation to research. Schwartz wants to protect research; so does Savary. But Schwartz claims that Savary's law ignores the subject. In fact the law is about education, and the true place of research is to be defined in a forthcoming decree. The decree, (a law defined by government independently of parliament) will define the rights and duties of university staff and will, say ministry officials, put great emphasis on the role of research. A recent, interim decree appeared to increase the teaching duties of professors and lecturers greatly, and made no distinction between the two categories; but, says the ministry, the full decree on rights and duties will allow a professor or lecturer to trade off teaching duties against research or administration time, and since these vary among categories there will be an effective variation of teaching load. The real question now to be resolved, says the ministry, is who will decide, and by what means, who can trade teaching for research and by how much.

Robert Walgate

India in Antarctica

Science — and politics — on ice

INDIA'S Antarctic ambitions continue to prosper. On 15 September, P.K. Basu, Secretary of the Indian Ministry of Mines, told the Central Geological Programming Board in Calcutta that a third scientific expedition to Antarctica is scheduled for 1983–84. Like the previous expeditions, it will carry out multidisciplinary scientific research and will again include geoscientists of the Geological Survey of India.

India's first foray to Antarctica began at Goa in December 1981 (see *Nature* 295, 640; 1982). The *Polar Circle*, an ice-breaker chartered from Norway, took the 21-man Indian team led by Dr S.Z. Qasim of the Department of Ocean Development to the Antarctic, where it landed on 9 January 1982 in the sector claimed by Norway. Logistical problems reduced the length of the stay from the planned 25 to 10 days, but a wide range of scientific observations was carried out. The expedition returned in February 1982, when Mrs Indira Gandhi, the Prime Minister, told the *Lok Sabha* (Lower House in Delhi) that "the main objective was to study the meteorology and other conditions of Antarctica, which are believed to control the monsoons" and to influence the climate of the Indian Ocean region.

To this end, an unmanned solar-powered weather station, *Dakshin Gangotri* (Southern Ganges), was established at 70° 45'S 11° 38'E in order to provide a continuous record over one year on a cassette to be retrieved by a second expedition. A further objective was to test the suitability of Indian equipment at sub-zero temperatures, while a significant discovery along the way was a sea-mount (named inevitably as Indira Mount) at 53° 22'S 48° 03'E, where it extended a line of sea-mounts already reported by the Soviet Union.

The second and much larger expedition of 28 scientists followed, and a third is planned, to prepare the ground for a permanent manned Indian scientific station to be established in 1985.

Clearly these expeditions reflect a natural desire to learn more about the relatively unknown Antarctic continent, which is thought to interrelate with various aspects of the Indian climate, geology and so on. In this respect, India is complementing the work of other nations, and particularly of the Scientific Committee on Antarctic Research (SCAR); in fact, the Antarctic Treaty of 1959 was designed largely "to promote international co-operation in scientific investigation in Antarctica" through the removal of political and other obstacles to research.

A further motive is the prestige attached to Antarctic research. Mrs Gandhi, whose personal enthusiasm played a major part in mounting the Antarctic project, implicitly admitted as much when she told the *Lok*

Sabha in February 1982 that the expeditions offered "one more proof, if such be needed, that Indian scientists and technologists have the capability to undertake the most hazardous and complex tasks . . . In undertaking this advanced work India has now joined a select band of countries".

Inevitably, however, other motives have been imputed to India, partly because, in Antarctica, it is difficult to disentangle scientific from political and economic considerations. Thus scientists have sometimes been regarded as political instruments, in the same way that increases in support for scientific research by the British Antarctic Survey have appeared to some to be more a political consequence of the Anglo-Argentine dispute over the Falklands than a result of a cool appraisal of the scientific possibilities.

In the past two decades, India has been the most articulate critic of the exclusivity of the Antarctic Treaty. India has argued that Antarctica should be treated as the common heritage of mankind, so that all countries should have equal rights to share in Antarctic decision-making and resources. India has thus been seen as a threat to the survival of the Antarctic Treaty system. During 1982–83, this threat seemed to increase not only because of India's two expeditions to Antarctica but also because of the interest shown by developing countries in a United Nations-based alternative to the Antarctic Treaty. Over the past year, for example, Dr Mahathir, the Prime Minister of Malaysia, has advocated the creation of a new Antarctic regime in the UN General Assembly (on 29 September 1982) and the Non-Aligned Summit Meeting in New Delhi (8 March 1983).

In the face of pressure from the non-aligned movement the UN General Committee agreed on 21 September 1983 to place Antarctica on the agenda of the current session of the General Assembly. Hitherto, the United Nations has steered clear, or rather has been steered clear, of Antarctica, but events there and at the Non-Aligned Summit suggest that Antarctica may become yet another focus for North–South controversy.

In this context, the intentions of India, one of the leading members of the non-aligned movement, proved a major pre-occupation of the Antarctic Treaty powers until, in August 1983, India astounded most observers by its accession to the Antarctic Treaty. (Accession entails acceptance of the treaty's principles.) India also applied for consultative party status — the right to participate in decision-making which is open only to those countries "active" in Antarctic research — and its application was approved by a special meeting of the consultative parties held in

Canberra on 12 September 1983. An application from Brazil was approved at the same time. So India is now part of the treaty system that it has heavily criticized in the past. The two expeditions can be seen in retrospect as attempts to satisfy criteria for membership of the Antarctic club.

Why has India confounded the experts, and apparently caused a serious rift in the anti-Antarctic Treaty camp? While disclaiming any territorial ambitions, India has been reticent in referring to the wider political aspects of its Antarctic policy. In 1982 Dr Qasim told *The Tribune* (published in Chandigarh) that "India considers Antarctica to be the common heritage of mankind and not a preserve of a few nations". It is clear that the Indian Government believes that a seat at the consultative party sessions will provide a better opportunity to influence developments in the southern continent, particularly in promoting the "common heritage" concept.

Qasim also conceded a more selfish motive when he said that India had ensured that it would not be left behind in any international race to exploit the hidden resources of Antarctica. The potential marine resources are viewed with particular interest in India, and it is probably significant that the Department of Ocean Development has been the main agency behind the Antarctic expeditions.

Scientific research has proved the chief beneficiary of the Antarctic Treaty system, and science will continue to benefit as long as the treaty survives. It is therefore ironic that, at the very time when the treaty system has been reinforced by recent additions to its ranks — 1982–83 saw not only the admission of Brazil and India as consultative parties but also the accession of Spain and China — it should face its most serious test yet. Internally, there have been difficulties surrounding the Antarctic mineral regime negotiations, which were begun in Wellington in January 1983, and continued at Bonn last July, while externally there is the Malaysian campaign to replace the treaty by a regime led by the United Nations, perhaps modelled on the International Sea-Bed Authority. Indeed, many countries feel that the principles underlying the Law of the Sea should be extended to Antarctica.

Scientists need to watch future developments closely, for any United Nations intervention in Antarctic affairs will be oriented towards resources rather than science. The recent enhancement of the Antarctic Treaty system — and the fact that it may soon be further strengthened by the addition as consultative parties of China, Spain and East Germany — encourages the view that the treaty will survive, so that any intervention by the United Nations will occur only within the parameters established by the 1959 treaty. If so, Antarctica may remain the "continent for science" advocated by Sir Vivian Fuchs in 1973.

Peter J. Beck

British universities

Survival by questionnaire

SIR Peter Swinnerton-Dyer, the new chairman of the University Grants Committee, has set out to alter the committee's reputation for secretiveness. Encouraged by Sir Keith Joseph, Secretary of State for Education and Science, to conduct an "open and wide-ranging" debate on the future of the British university system over the next ten years, the committee has published in full a circular letter sent to universities.

The letter takes the form of an all-embracing series of Catch-22 questions on what changes the universities would like to see — or be able to tolerate. Although the questions are addressed to vice-chancellors, Sir Peter expects many of them to be answered by particular groups within universities, to avoid delay and "answers with the consistency of babyfood".

On the central question of how resource per student will change, universities have been asked to consider a number of options specified by Sir Keith. They range from level funding, in real terms, to a 2 per cent annual decrease. In response to a request from Sir Peter's predecessor to find places for an extra 5,000 students, next year and in the year following, at no extra cost, universities have submitted proposals for 2,500 more science students and 1,000 more humanities students for each of the years in question. But Sir Peter was unable earlier this week to give any clear idea of how resources might change in future, and observed that the Department of Education and Science is unlikely to have a master plan that it is working towards. Universities' responses may in part determine future policy.

Despite the long term uncertainty, the fact that resource per student will certainly not increase, together with the fact that student demand seems certain to decrease later this decade, must imply a substantial contraction of the university system. If the present balance between universities and the public sector is maintained, the universities must shrink by 15–20 per cent during the early 1990s. Against this background, universities are asked for their views on changes in organization and subject balance.

Sir Peter refused to rule out the possibility of closures, but observed that there would be serious constitutional difficulties in such a course. To revoke a Royal Charter is without recent precedent (although this was apparently done by James II of England before he was deposed in 1688). But Sir Peter pointed out that several universities are now as small as is economic — about 4,000 students for a university offering a full range of courses. The most likely scenario is that some universities might merge with polytechnics or other colleges, and cease to be supported mainly by the University Grants Committee.

Whatever happens, there will certainly

be a blurring, if not a disappearance, of the "binary line" between universities and other institutions of advanced further education. And any plans for the universities must be reconciled with plans now being drawn up for the public sector.

The government is also keen to encourage universities to find other sources of funds than the public purse, and financial links with industrial companies seem likely to increase. One of the questions now asked of the universities is disarmingly frank: do you think that the dual support system can survive, and would you wish it to do so? But Sir Peter conceded that there may be difficulties in asking some universities to find all their support from industry, although he observed that there are excellent universities in Japan supported in this way.

One major problem taxing the committee is the age distribution of academic staff. As most universities coped with the 1981 cuts by encouraging early retirement, the present distribution is very unbalanced. The replacement rate now in prospect is below that considered desirable by the Department of Education and Science a year ago. One possibility is that the "New Blood" scheme for recruiting extra academic staff might be extended beyond the three years planned. But some researchers are fearful that the scheme could be used as an instrument to force unwelcome change: after the first year of an appointment under the scheme, allowances are subsumed within universities' recurrent grant. The question of security of tenure is also raised in this context. The inviolability of tenure has still not been fully tested, but the question cannot be put off indefinitely.

Universities are also asked to consider whether the research component of their recurrent grant should be "earmarked". This controversial proposal has been made several times recently, and is favoured by some researchers. But, although the University Grants Committee is exercising "positive self-restraint" in not pre-judging issues before universities have replied, it was made clear that the committee sees unusual and difficult problems in this course, chief of which is the difficulty of accurately costing the time of someone who is both a teacher and a researcher.

Many of the questions could be fairly described as speculative — for example, how would universities react to a replacement of the present system of sixth-form education by a broader alternative? And universities are encouraged to add any further comments of their own on whatever subject they wish. Answers are requested by 31 March 1984, and will be marked by the following October. Results (not, apparently, graded) will then be sent to the Secretary of State. But this will be only the first stage of the debate.

Tim Beardsley

Semiconductors

Missing link from Inmos?

INMOS International PLC, the government-backed British semiconductor manufacturer, last week gave a description of a new type of microprocessor which, it claims, will become a building block for "fifth generation" computers.

The device, to be known as the transputer, was originated by Mr Iann Barron, a director of Inmos. The transputer includes on a single chip a processor, 4 kbytes of memory and communications links. The first member of the family will be a 32-bit device running at 10 million instructions per second, claimed to be eight times faster than any existing or announced device. A 16-bit version is also planned.

The transputer will operate with a reduced instruction set optimized for execution of high-level languages. The impressive specification has been achieved partly by increased communication speeds possible when memory and processor are combined on the same chip. But the feature that is likely to secure a long-term future for the transputer is that it is designed for integration into arrays of many transputers operating concurrently. The transputer will be able to interface with existing industry standard devices and, most importantly, with other transputers, via special "Inmos links". These links will allow an array of transputers to operate as an "intelligent community". Inmos has developed a special language for use with the transputer which reflects this feature. The language, called Occam, allows problems to be broken down into separately-running subunits so that the processing speeds likely to be needed in so-called intelligent machines can be achieved.

Inmos is still looking for further sources of capital. In line with recommendations from the government, the British Technology Group has indicated that it will not be able to shore up the company much longer. Inmos made a loss of more than £20 million last year, but hopes to reach profitability next year if market conditions improve. Indeed, if the transputer is as successful as a hopeful spokesman for Inmos expects, there may be a problem in meeting demand.

Tim Beardsley

Correction

In the leading article entitled "Reagan flirts with pie in the sky" (*Nature* 3 November, p.1), the proposal is to develop satellites mounting enough weapons to destroy 10,000 re-entry vehicles (not 1,000 as stated), and the cost of the first stage of the work would be \$95,000 million (not £95,000 million).

The first News story (p.3) concerns NASA's 1985 budget submission, not that for 1984. □

UK space research

Earth scientists make their bid

BRITISH oceanographers, climatologists and solid-Earth scientists set their sights last week on a ten-year programme of satellite-based radar altimetry. At a meeting organized by the Royal Astronomical Society, more than a hundred geophysicists made common cause by advocating that the Science and Engineering Research Council (SERC) and the Natural Environment Research Council (NERC) should spend £18 million for the development of instruments that would be carried by various of the remote sensing satellites now being planned, the Canadian Radarsat for example.

The proposal is, however, politically sensitive because such an extra claim on research council budgets may be resisted by other dependants of the research councils, particularly the astronomical community.

The benefits of radar altimetry were loudly sung last week. Thus Professor J. Woods of the University of Kiel explained that, while flow patterns within major oceans and seas are well known, uncertainty persists about even the largest scale flows between such basins. Yet climatic modellers concerned with the effects of atmospheric carbon dioxide need to know the oceanic fluid and heat circulation which cannot now be supplied by oceanographers. The only alternative to satellite altimetry, according to Professor Woods, would be a massive sea monitoring exercise.

Altimetry may also be used to monitor the elevation and extent of ice sheets on land and sea, of obvious significance in the

relatively inaccessible Arctic and Antarctic ice fields. Again, the extent of large ice sheets is climatologically significant, given their ability to affect ocean temperature, shallow and deep circulation patterns and also the amount of sunlight reflected back into space. For the solid Earth, satellite altimetry provides the detailed shape of the Earth's gravitational field which, for example, can be related to the processes of convection in the Earth's mantle which are thought to drive plate tectonics.

For such a project, accurate positioning of satellites by laser and radio Doppler tracking would be essential. The accuracy of altimetry is expected to be 10 cm or so over a range of 1,000 km, and laser tracking can measure heights to comparable or greater accuracy. In comparison, the elevation amplitudes of oceanographic phenomena range between 10 cm for currents on continental shelves, 30 cm for changes in heat storage (over periods of months) and 1 m for tides and the Gulf Stream. Accuracy to 10 cm is required for useful analysis of the gravitational field.

On the political aspects of the project, the idea is that the Advisory Board for the Research Councils (ABRC) should make a bid for an extra £18 million from the Department of Education and Science over ten years to cover hardware and software development. Suppliers of the project argue that it should give the United Kingdom a lead in the technology and interpretation of precision altimetry. But the budgets of SERC and NERC would need to be maintained to provide back-up

Harvard to pay up record amount

Washington

HARVARD University has agreed to repay \$4.6 million improperly charged to federal research grants. The agreement settles a six-year investigation of Harvard's management of research grants by auditors from the Department of Health and Human Services (HHS, the parent agency of the National Institutes of Health), which found that funds were routinely transferred from one project to another in order to cover cost overruns and to ensure that all available grant money was spent. Audits of the school of public health and the medical school found evidence of \$3.8 million in improper charges as well as tens of millions of dollars in charges that could not be verified because of inadequate documentation.

When the audit of the medical school was released last year, Harvard officials said they had no intention of repaying anything more than a token amount (see *Nature* 300, 3; 1982). Explaining the change of heart, Harvard's vice-president

for finance, Thomas O'Brien, said that although an audit commissioned for Harvard for fiscal year 1978 showed an "error rate" of less than 2 per cent, which the university felt was acceptable, the government "could accept no error rate at all". O'Brien said that given the cost of litigation, Harvard agreed to pay up.

Under the agreement, the university will repay \$1 million in cash immediately. The remaining \$3.6 million will come from adjustments in overhead payments on research grants to the university over the next three years. According to HHS, the final figure was arrived at by extrapolating the government's audit findings, which cover 1975 to 1977, up to 1983. Harvard has also agreed to improve its accounting procedures. HSS auditors say the \$4.6 million is the largest single repayment from a university, although in relative terms it amounts to less than 1 per cent of the federal research funds received by Harvard during the eight years from 1975 to 1983.

Stephen Budiansky

research and data analysis.

Some argued last week that, at a time when both research councils are under financial stress, it would be better for geophysicists to compete more strenuously with other disciplines for existing research council funds. The current structure of the councils means, for example, that the radar altimetry project is in direct competition

with big spenders such as the X-ray astronomers.

Sir Robert Boyd of the Mullard Space Science Laboratory, himself an X-ray astronomer, argued that in times of shortage, a greater proportion of the money available should go towards such projects as radar altimetry which promise important scientific information and also

provide more direct industrial and technological spin-off than some of the more glamorous astronomical projects. However, calls from some speakers for more open warfare within the grant-awarding committees were dismissed by others as being potentially disastrous for science as a whole.

Philip Campbell

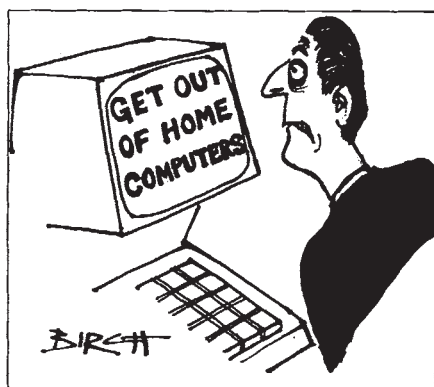
US personal computers

Shakeout comes at last

Washington

THE long-forecast shake-up in the overcrowded personal computer industry arrived with a vengeance last week. Texas Instruments announced that it was pulling out of the home computer market and ending production of its 99/4A home computer. A day later, the International Business Machines Corporation (IBM) unveiled its Personal Computer Junior (PCjr), the product previously codenamed the Peanut and the cause of turbulent trading on the stock market for the past two months.

IBM's PCjr, which will be on sale early next year, will cost \$700 and is designed to complement the company's more powerful Personal Computer, which has already captured more than a quarter of the market for machines in the \$1,000 to \$5,000 price range. The basic machine will have 64



kbytes of memory but a 128 kbyte version with a disk drive will be available for \$1,300. A key selling-point will be the fact that the PCjr is compatible with the Personal Computer, so that buyers who already use the Personal Computer in their offices will be able to use many of the same programs at home.

The fact that the PCjr has drawn mixed reviews from trade experts hardly matters. IBM's name and its growing market dominance should ensure that the new machine will sell well and that software publishers will increase their efforts to make their programs IBM-compatible. Since introducing its Personal Computer two years ago, IBM has already forced Apple, previously the leader in the small computer market, very much into second place. Apple's market share has fallen from 29 per cent in 1981 to 17 per cent this year.

With the success of IBM virtually

guaranteed, other competitors are beginning to lead the field. The Osborne Computer Corporation, one of the pioneers of small computers, declared bankruptcy last month. And Texas Instruments' decision to leave the home computer market is a sign that even large corporations have become disenchanted. The company's withdrawal sent the price of its shares soaring from \$22.75 to \$124.50, but added little to the value of the shares of its principal rivals in the small computer market, Commodore and Atari (a subsidiary of the Warner Corporation).

Texas Instruments' 99/4A was a popular machine but one that became a victim of the sharp price-cutting war that has characterized the home computer industry. By concentrating on its strengths in semiconductor manufacture and defence electronics, the company hopes to compensate for a \$200 million loss this year which it attributes to poor sales of home computers.

A reduction in the number of companies manufacturing home computers may be good news for consumers. Although less competition may allow prices to rise, standardization of the industry around the hardware of several major companies will

enable software publishers to write programs for markets that have until now been fragmented. Educational software, for example, has been relatively neglected. In congressional hearings recently, Secretary of Education Terrel Bell said the incompatibility of the computers bought by schools had hindered the development of high quality teaching programs.

The drive towards standardization is not confined to the United States, however. In Japan, where the absence of standardized software has been the bane of the home computer industry, five companies — Hitachi, Matsushita, Mitsubishi, Sanyo and Toshiba — have introduced new computers standardized around the MSX operating system produced by a US company, the Microsoft Corporation of Bellevue, Washington.

William Gates, Microsoft's chairman, said the arrangement would open the Japanese home computer market to software companies in the United States and Europe. The US Department of Commerce, however, is concerned about the corollary: it will also open the United States home computer market to Japanese companies. Clyde Prestowitz, an official of the Department of Commerce, said both Apple and the Commodore Corporation had complained about the implications of the Microsoft deal.

Peter David

Nature index of biotechnology stocks

12-Month high	12-Month low	Company	Close previous month	Close 28 Oct.	Change
23 1/4	10 3/4	Biogen (Switzerland)	14 1/4	12	-2 1/4
6 1/4	2	Bio-Logicals (Canada)	2 1/2	2 1/8	- 3/8
16 1/8	7 1/4	Bio-Response (USA)	14 7/8	11 3/4	-3 1/8
19	11 5/8	Cetus (USA)	12 7/8	11 7/8	-1 1/4
15 1/2	8 7/8	Collaborative Research (USA)	11 1/8	9 1/2	-1 5/8
39 7/8	15	Damon (USA)	25 1/2	20 1/8	-5 3/8
34 1/4	16 3/8	Enzo-Biochem (USA)	25 1/2	24 1/4	-1 1/4
18 7/8	8 5/8	Flow General (USA)	8 3/4	9 1/4	+ 1/2
49 3/4	25 7/8	Genentech (USA)	35 1/2	26 1/2	-9
17 3/4	8 3/4	Genetic Systems (USA)	11 1/2	9 1/4	-2 1/4
23 1/4	12 7/8	Genex (USA)	18 3/4	15 1/2	-3 1/4
31	19 1/2	Hybritech (USA)	19 3/4	22 1/2	+ 2 3/4
22 1/4	12 1/4	Molecular Genetics (USA)	17	13 1/4	-3 3/4
23 1/4	13	Monoclonal Antibodies (USA)	14	14 1/4	+ 1/4
73 1/2	42	Novo Industri A/S (Denmark)	67	66	-1
30 1/8	13 1/2	Pharmacia (Sweden)	26 1/2	25 1/8	-1 3/8

Closing prices are for the last Friday of the month. For over-the-counter stocks, bid price is quoted; for stocks on the American and New York exchanges, the transaction price. *Nature's* weighted index of biotechnology stocks stood at 203 on 28 October, compared with 221 a month earlier. Data from E.F. Hutton, Inc.

Animal do have rights

SIR — As an occasional contributor¹ to the burgeoning philosophical literature on animal rights, I find myself dismayed by both the content and tone of your editorial, "Animal rights nonsense" (*Nature* 13 October, p.562). The editorial appropriately bemoans the absence of "the voice of reason from the middle ground. . ." but unfortunately exemplifies this lacuna in its snide treatment of its subject.

No philosopher has seriously proposed granting animals ethical or legal rights that are fully equivalent to those possessed by normal human beings. Discussions of the "responsibilities" of animals, while perhaps clever in a sophomoric sort of way, simply fail to engage the moral question at issue. The question of the treatment of animals in experimentation, agriculture or in any other setting concerns not the status of animals as moral agents but the legitimacy of their status as objects of moral concern and consideration. The philosophical question fuelling concern with animal rights is whether animals have any moral standing at all, not whether we should place them on juries or place them on trial for their misdeeds as responsible and culpable moral agents.

If you really want to seek the middle ground in the animal rights debate it would be more to the point to address the following philosophical conundrums; what properties confer moral worth on a creature, be it human or animal, what should we do when the needs of human beings come into conflict with the interests of animals and what principles of morality should govern the conflicts that inevitably arise among and between both animals and humans in a world of scarcity, uncertainty and risk?

The last question is the one that philosophical proponents of animal rights have failed adequately to address. Like their conceptually muddled forerunners in the abortion debate, many animal rightists seem content with the demonstration that animals have some rights. This being achieved through the invocation of various criteria such as sentience or purposiveness, proponents of animal rights think matters closed — if animals have rights then humans must respect them.

However, the demonstration of a basis for the moral standing of animals hardly ends moral discussion. Indeed, one can only begin an ethical examination of the human treatment of animals on the basis of the fact that animals have some rights worth talking about! The nub of the matter in need of resolution is what to do when rights or, as I much prefer, interests come into conflict. On this thorny point, animal rightists and, sadly, it would appear from the intemperate editorial, the editors of *Nature* have found common ground — neither has yet been able to say anything useful about what moral principles policy

makers, legislators and the scientific community ought to use in resolving the unavoidable conflicts of interest that sometimes exist between animals and humans.

ARTHUR L. CAPLAN

*The Hastings Center,
New York 10706, USA*

1. Caplan, A.L. *Ann. NY Acad. Sci.*, 159-169 (1983).

Sutton's role

SIR — In the otherwise excellent article (*Nature* 13 October, p.575) regarding the well-deserved prize and recognition of Barbara McClintock, I found erroneous the statement, (her) . . . "observation, which may be regarded as the foundation of cytogenetics". I believe that Walter S. Sutton is usually credited with having established this foundation in an article entitled "The Chromosomes in Heredity"¹. His beautifully reasoned work is cited often in this context; for example in *Classic Papers in Genetics*²: "Today the two fields are permanently linked . . . and Sutton's paper can be considered the beginning of the field called cytogenetics". And J. D. Watson wrote "Sutton's paper . . . brought together for the first time the independent disciplines of genetics (the study of breeding experiments) and cytology (the study of cell structure)"³.

Interestingly, as you mentioned, McClintock's first studies were based on early work done at the University of Missouri with corn while Sutton's first work, not very far away at the University of Kansas, was with the giant chromosomes of the grasshopper.

SANTIAGO GRISOLIA
*Instituto de Investigaciones Citológicas
de la Caja de Ahorros de Valencia,
Valencia-10, Spain*

1. Sutton, W.S. *Biol. Bull.* 4, 231-251 (1903).
2. Peters, J.A. (ed.) *Classical Papers in Genetics*, 27 (Prentice-Hall, Englewood Cliffs, 1959).
3. Watson, J.D. *Molecular Biology of the Gene* 2nd edn, 18 (Benjamin, New York, 1970).

• Dr Grisolia is right — Editor, *Nature*.

Schizophrenic terms

SIR — A footnote to the comment by Iversen *et al.* on schizophrenia and the claimed "causal" role for dopamine or noradrenaline¹. The caution with which the authors disclaim "cause" in favour of some sort of correlation and a role of neurotransmitter in the mediation of the therapeutic effect of drugs is helpful. However, in the interests of getting our thinking straight about the complex phenomenology of schizophrenia, can they — and your other correspondents — be encouraged to avoid the use of the term "schizophrenic brain" to refer to post-mortem samples of brain tissue derived from patients diagnosed as schizophrenic? Schizophrenia is a term used to describe a particular pattern of behaviour and

expressed feelings in an entire human individual in his or her multiple interactive relationships with fellow humans. To reduce is — at best — to assume what the biomedical model of schizophrenia is trying to prove. At worst, it is a step down that same reductionist path that leads to "genes" for depression, or indoctrination or homosexuality — or even "selfish" DNA. Anthropomorphizing non-human animals is bad enough; anthropomorphizing tissues or molecules is just sloppy. Pleading that it is only a shorthand, and we all *really* know what we mean don't we, is no defence, granted the poor history of clear thinking in this terrain^{2,3}.

STEVEN ROSE

*Biology Department, Open University,
Milton Keynes MK7 6AA, UK*

1. Iversen, L.L., Reynolds, G.P. & Snyder, S.H. *Nature* 305 577 (1983).
2. Rose, S.P.R. in *Biological Markers and Mental Disorder* (W.H.O., in the press).
3. Rose, S.P.R., Lewontin, R.C. & Kamin, L. *Not in Our Genes* (Penguin, Harmondsworth, in the press).

The "wine-dark" sea

SIR — R.H. Wright and R.E.D. Cattle ingeniously suggest (*Nature* 303, 568; 1983) that Homer called the sea "wine-dark" because his wine, when mixed with water with a high pH factor, turned blue. However, we know from other references that his wine was red (*Odyssey* 5.165; 9.163), black (5.265; 9.196, 346) or dusky (*Iliad* 1.462, 4.259), hence probably like modern Mavrodaphni wine. He also describes cattle as "wine-dark" (*Odyssey* 13.32; *Iliad* 13.703) — surely not blue!

The October issue of *Greece and Rome* provides another possibility. High dust content in the atmosphere gives a dark red sunset. When this sunset is reflected in a dark, oily sea, it gives a colour and texture very close to that of Mavrodaphni. I first observed this phenomenon in a tidal estuary of the Damariscotta River in Maine, under the dust cloud from an eruption of Mt St Helens on the west coast.

Further examination of the references to "wine-dark sea" shows that the phrase is normally used of weather conditions at dusk. All but two formulaic uses of the phrase without context can be explained as references to a dark red sea (once under a dark storm cloud, once as a simile for the colours under a storm cloud, *Iliad* 5.770-2). Except when the colour is caused by storm, it would be a good weather sign for ships steering at night by star navigation. Dusty skies are a reliable sign of good weather behind. Ancient ships preferred to sail by day, hugging the coast, but we assume that for long voyages over open ocean they preferred nighttime star navigation, on which there was an ancient manual, and set out under favourable weather signs such as a "wine-dark sea".

R. RUTHERFURD-DYER

*58 N. Prospect,
Amherst, Massachusetts 01002, USA*

Can observation prevent decay?

Failure so far to detect proton decay is not explained by the suggestion that intra-nuclear collisions prevent it. But there is a quantum version of Zeno's paradox to be resolved.

ALTHOUGH quantum mechanics is more than half a century old, arguments persist about the most elementary features of the theory, even the question of what precisely are the consequences of a measurement of some physical attribute of a quantum system. But did not Heisenberg settle the matter once and for all with his statement of the uncertainty principle, the notion that the uncertainty in the measurement of a pair of canonically conjugate dynamical variables (such as position and momentum, or energy and time) must be reciprocally related in the sense that the product of the two uncertainties is bounded below by Planck's constant? Unfortunately not, as can be told from a row about the lifetime of the proton that has blown up in the pages of *Physical Review Letters*.

To be sure, it follows from the uncertainty principle that a precise measurement of the momentum of, say, a particle leaves its position essentially undetermined. One of the reasons for the resurgence of interest in the past few years in the theory of quantum measurement is the ambition to extract the utmost information from equipment designed for the measurement of signals virtually buried in receiver noise, the detection of gravitational waves for example. One result is the growth in the past few years of the minor industry concerned with what are known as "quantum non-destruction" measurements. Now, however, an even more elementary issue has bubbled up — that of whether it is possible to prevent some phenomenon (say the radioactive decay of an unstable nucleus) from taking place at all by keeping the system (that is, the nucleus) under continuous observation.

The argument goes like this. For many quantum systems, for example atoms with less energy than that needed to detach the least strongly bound electron into the energy continuum, most measurements of dynamical quantities yield only one of a set of discrete values, the eigenvalues of the appropriate Schrödinger equation or its equivalent in the Heisenberg matrix formalism. Everybody knows that. So what happens if one measurement is quickly repeated? The effect of the first measurement is not merely to yield as a numerical result one of the possible eigenvalues but also to ensure that the system as a whole is driven into the corresponding eigenstate. So if the

measurement is quickly repeated, the result will be exactly the same numerical value.

In the mid-twenties, while the Schrödinger and Heisenberg schools were developing techniques for calculating the allowable eigenvalues for particular quantum systems, Bohr's Copenhagen school was working out the formal language in which these questions are still discussed. At the outset, before any measurements have been made, the state of the system is literally undetermined, some indeterminate mixture of all possible states. The effect of the measurement is to resolve this uncertainty. The Copenhagen word for this process is that of the "preparation" of a system in a particular eigenstate. Once that has been done, the system will stay as it is unless influenced by extraneous forces (such as those occasioned by measurements of a conjugate variable). So if arrangements are made to make some measurement of the state of an unstable nucleus, and to repeat that observation frequently, should not the result be that the system persists in its prepared state and does not decay?

Earlier this year, L.P. Horwitz and E. Katznelson (*Phys. Rev. Lett.* **50**, 1184; 1983) advanced an argument very much like this to suggest that protons bound within nuclei should not decay at anything like the rate suggested by the now fashionable gauge theories of the strong nuclear interaction — and thus to explain the failure so far to detect such decay. What the two authors (from Tel Aviv University, Israel) suggested was that the intranuclear collisions with other nucleons to which any potentially unstable proton must be subjected would ensure its survival in its prepared state — that of a proton. For even though the collisions take place at random intervals, and although the outcome of these collisions is not self-consciously recorded, if a proton should at some instant have decayed (say into a π^0 meson and a positron, one of the more probable decay routes), the result would be that the dynamical behaviour of colliding nucleons would be transformed in such a way as to produce an excited state of the nucleus as a whole. Horwitz and Katznelson therefore argued six months ago that repeated internal collisions would be entirely equivalent to a sequence of measurements repeated at intervals of time comparable with the average time between nucleon collisions. For their temerity, they have brought down on their heads an impressive declaration of

dissent (*Phys. Rev. Lett.* **51**, 1599; 1983).

The argument is nevertheless not devoid of interest. The three refutations now published make a number of separate points, of which the most persuasive (nevertheless denied by Horwitz and Katznelson) is that a nucleus is a single entity described by a set of eigenstates whose properties are determined not merely by the constitution of the nucleus but also by the interactions between them. On this view, intranuclear collisions between nucleons are accounted for from the outset, with the consequence that their occurrence is irrelevant to the enumeration of the quantum states in which the nucleus exists, or between which transformations can recur. In passing, however, the counter-arguments have the interest of drawing attention to an often-forgotten feature of decay processes — that while the rate of the decay of a system is usually correctly taken to be a constant, at the outset of the existence of an unstable state there is a short period during which the likelihood of decay is proportional to the square of the elapsed time, which may be put at $A^2 t^2$, where A is a constant and t is elapsed time. Then the probability that the system will not have decayed is $(1 - A^2 t^2)$ and, if it is imagined that a large number of measurements, say N , is carried out at intervals of t , the probability that the system will not have decayed after an elapsed time of Nt is $(1 - A^2 t^2)^N$, which works out as $\exp(-AT^2/N)$ where T is the total elapsed time Nt and N is assumed large. Since N may be arbitrarily large, the chance of a system not having decayed after the passage of time T may be made arbitrarily close to unity.

Kevin Cahill of the University of New Mexico, one of the tormentors of Horwitz and Katznelson, points to this paradoxical conclusion before going on to argue that the paradox is unreal because the collisions do not constitute observations. But can the paradox be real? Coincidentally, but apparently without the knowledge of Horwitz and Katznelson and their critics, the question has been taken up by M. Bunge and A. J. Kalnay in *Il Nuovo Cimento* (**B77**, 10; 1983), who explicitly deal with the general suspicion that the paradox must be a fraud. Their conclusion is simple — the false assumption in the calculation summarized above is that measurements can be carried out instantaneously. In reality, this cannot be so — and the process of measurement can be as much a stimulus to transformation as to stasis.

John Maddox

Cell biology

Muscle *au naturel*

from Roger Craig

THE rapid-freeze deep-etch rotary-shadow technique of electron microscopy¹ has generated considerable excitement in cell biology circles. By means of it, the internal structure of cells can be studied *au naturel*, without marinating in glutaraldehyde, osmium and alcohol or cooking in epon or araldite. Following rapid freezing (by contact with a liquid helium-cooled copper block), the specimen is fractured and etched; thenceforth the only 'cooking' it receives before viewing in the electron microscope (EM) is a thin frosting of platinum and carbon evaporated on to the specimen at a temperature of -100°C . By avoiding the culinary procedures of conventional EM, rapid freezing should not only provide a picture of most cells closer to the *in vivo* state than usual, but it is particularly applicable to delicate fixation-sensitive specimens and to the study of transient phenomena, which can be arrested within 2 milliseconds. Thus, among many applications, the quick-freeze technique has provided dramatic three-dimensional views of the interior of intestinal epithelial cells and their microvilli² as well as a visual time course of synaptic vesicle exocytosis at frog neuromuscular junction³.

Although muscle is perhaps most widely appreciated in its cooked form, we are fortunate now to be able to savour it raw as well. Indeed muscle is as apt a tissue as any to be studied by rapid-freezing, since the rapid molecular motions of the myosin cross-bridges, which bring about contraction, remain inscrutable (despite intensive study using time-resolved X-ray methods)⁴ and quick-freezing gives hope of revealing them. Three papers in a recent issue of the *Journal of Molecular Biology* take us a step towards this goal. In two of these, Heuser and Cooke^{5,6} give us dramatically detailed three-dimensional views of insect flight muscle in the relaxed and rigor states, as well as a new look at an old structure, the actomyosin arrowhead complex, while in the third⁷ Heuser demonstrates how his method can now be applied to single molecules (as well as cells), amongst them myosin, by adsorbing them to tiny mica flakes before freezing.

The most comforting finding is that Heuser's unfixed freeze-etched specimens provide a picture of muscle filaments and molecules largely similar to those that have been obtained by conventional EM — the standard recipes painstakingly developed over the years have not been in vain. Deep-etched rotary-shadowed myosin molecules look very similar to those studied by the now widely used glycerol/mica rotary-shadow technique pioneered by Elliott and Offer⁸ and more recently modified⁹⁻¹¹. The

molecule is seen to consist of a long thin tail attached at one end to two pear-shaped heads. Actin filaments look rather different from the standard negatively stained image: the left-handed $\sim 5.9\text{-nm}$ pitch genetic helix is emphasized while the two long-pitch helices are less obvious. The filament diameter is as great as 9 nm (after correction for metal thickness) rather than the generally accepted $6\text{--}8\text{ nm}$ seen in sections or negatively stained specimens. These observations are consistent with a recently proposed model for the actin filament¹² and Heuser and Cooke suggest that the smaller diameter usually observed may be due to shrinkage and/or stain penetration into the actin subunits in conventionally contrasted actin.

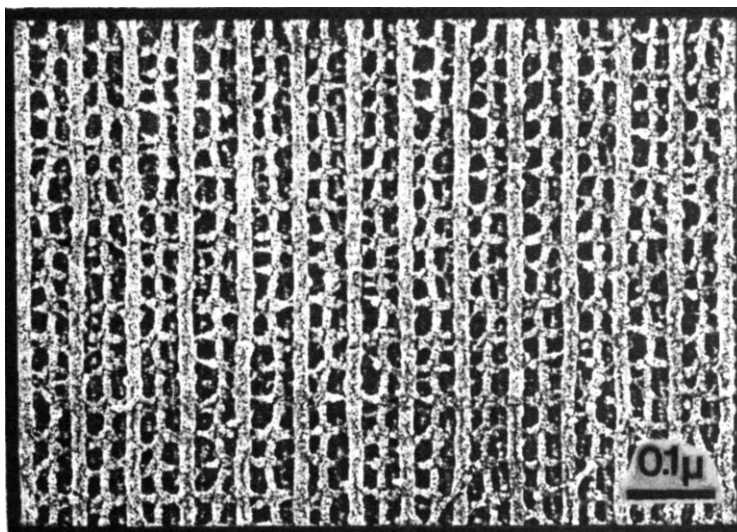
Actin filaments decorated with myosin subfragment-1 give the appearance of strings of polar arrowheads when observed by uranyl acetate negative staining¹³. When unfixed specimens are observed by Heuser's method, however, they appear simply as two-stranded ropes with no obvious polarity. This apparent discrepancy in fact reinforces the interpretation of the negatively stained images. The surface appearance of three-dimensional reconstructions of arrowheads, calculated from negatively stained images, displays much less polarity than the arrowheads themselves (for example, see ref. 14), and it is such a surface view that Heuser's method provides. It is superposition effects that occur in negatively stained specimens as viewed in projection in the EM that largely give rise to the polar arrowhead appearance. Heuser and Cooke make a further interesting observation — that decorated actin that is fixed and stained before rapid-freezing does display a subtle polarity not seen without fixation.

They conclude that pretreatment does not create the polarity but conserves a polarity which is lost in quick-frozen specimens that are not pretreated. Thus even with specimens prepared by the near-to-ideal quick-freeze method, the results must still be viewed circumspectly.

The beautiful and detailed longitudinal views of rigor insect flight muscle prepared by Heuser's method show the myosin crossbridges linking thick and thin filaments in dramatic three-dimensional relief (see accompanying micrograph). The results directly confirm Reedy's¹⁵ detailed observations and analysis of double-chevrons and flared Xs which were based on standard thin sections. Surprisingly, however, the 45° tilted state of the crossbridges seen by Reedy and generally assumed to be diagnostic of the rigor state, is not seen in unfixed specimens, but only when muscles are fixed (under the common, mildly acidic conditions) before freezing. In unfixed specimens the crossbridge angle is close to 90° : Heuser and Cooke suggest that the 45° angle is an artefact generated by shrinkage of the filament lattice that occurs with normal preparative methods. If this is so, it still appears that with the *in vivo* lattice spacing (determined by X-ray diffraction), a considerable tilt ($50\text{--}60^{\circ}$ from the filament axis⁵) is to be expected and that the $80\text{--}90^{\circ}$ seen in unfixed frozen specimens may be due to some unnatural lattice expansion. In support of this, modelling of the X-ray diffraction patterns of intact rigor muscles appears to demand a considerable cross-bridge tilt¹⁶⁻¹⁸.

Lastly, Heuser turns his attention to the appearance of thick filaments in relaxed muscle. In accord with conventional wisdom, few of the crossbridges are attached to the thin filaments, and the thick filaments are wider than in rigor, and irregularly dotted with bumps. Surprisingly, however, there is no 14.5-nm repeat of projections. Heuser tentatively attributes this to large thermal motions of the crossbridges, which are trapped by rapid

Actomyosin layers in unfixed rigor muscle.



freezing, implying that the 14.5-nm repeat seen by other methods is due to the underlying thick filament shaft. X-ray data show, however, that the 14.5-nm reflection in relaxed muscles arises largely from structures at high radius, which can only be the crossbridges^{19,20}; and conventional negative staining methods have preserved the 14.5-nm repeat in a variety of thick filaments, likewise showing that it arises from structures projecting from the backbone (refs 21–23, and my own unpublished observations of insect thick filaments). The order of the native crossbridge array in relaxed muscle has apparently been lost in Heuser's pictures (presumably during

fracturing or etching), suggesting that here too caution in interpretation of quick-freeze images is the order of the day.

Heuser's *nouvelle cuisine* electron microscopy provides a view of cell structure in as close to native state as has yet been achieved. The stage is now set for a study of transient crossbridge events in contracting muscle. If the results must still be taken with a grain of salt, Heuser's studies provide us with an appetizing flavour of things to come. □

Roger Craig is an Assistant Professor in the Department of Anatomy, University of Massachusetts Medical School, Worcester, Massachusetts 01605.

1. Heuser, J. *Trends biochem. Sci.* **6**, 64 (1981).
2. Hirokawa, N., Tilney, L.G., Fujiwara, K. & Heuser, J.E. *J. Cell Biol.* **94**, 425 (1982).
3. Heuser, J.E. & Reese, T.S. *J. Cell Biol.* **88**, 564 (1981).
4. Huxley, H.E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2297 (1981).
5. Heuser, J.E. & Cooke, R. *J. molec. Biol.* **169**, 97 (1983).
6. Heuser, J.E. *J. molec. Biol.* **169**, 123 (1983).
7. Heuser, J. *J. molec. Biol.* **169**, 155 (1983).
8. Elliott, A. & Offer, G. *J. molec. Biol.* **123**, 505 (1978).
9. Fowler, W.E. & Erickson, H.P. *J. molec. Biol.* **134**, 241 (1979).
10. Shotton, D.M., Burke, B.E. & Branton, D. *J. molec. Biol.* **131**, 303 (1979).
11. Tyler, J.M. & Branton, D. *J. ultrastruct. Res.* **71**, 95 (1980).
12. Egelman, E.H. & DeRosier, D.J. *J. molec. Biol.* **166**, 623 (1983).
13. Huxley, H.E. *J. molec. Biol.* **7**, 281 (1963).
14. Wakabayashi, T. & Toyoshima, C. *J. Biochem. (Tokyo)* **90**, 683 (1981).
15. Reedy, M.K. *J. molec. Biol.* **31**, 155 (1968).
16. Wray, J., Vibert, P. & Cohen, C. *J. molec. Biol.* **124**, 501 (1978).
17. Holmes, K.C., Tregear, R.T. & Barrington Leigh, J. *Proc. R. Soc. B* **307**, 13 (1980).
18. Offer, G., Couch, J., O'Brien, E. & Elliott, A. *J. molec. Biol.* **151**, 663 (1981).
19. Huxley, H.E. & Brown, W. *J. molec. Biol.* **30**, 383 (1967).
20. Wray, J.S., Vibert, P.J. & Cohen, C. *Nature* **257**, 561 (1975).
21. Reedy, M.K. & Garrett, W.E. in *Insect Flight Muscle* (ed. Tregear, R.T.) 115–136 (Elsevier, Amsterdam, 1977).
22. Stewart, M., Kensler, R.W. & Levine, R.J.C. *J. molec. Biol.* **153**, 781 (1981).
23. Vibert, P. & Craig, R. *J. molec. Biol.* **165**, 303 (1983).

Molecular biology

A straight LINE story

from John Rogers

SEVERAL per cent of the human or mouse genome consists of repeated units several kilobases long, interspersed throughout the genome. These 'long interspersed elements' or LINEs^{1,2} are obvious to anyone who performs gel electrophoresis on genomic DNA cut with a restriction enzyme; they stand out from the complex mixture as characteristic restriction fragments. But what are they? DNA sequencing has now shown that human and mouse each have a single major LINE family, and that they are related to each other. And the details of the sequences suggest that most, if not all, copies are inserts of reverse transcripts.

It is already widely believed that the shorter interspersed sequences, of which human *Alu* is the prototype, are insertions^{3,4} produced by reverse transcription. The hallmarks of such inserts are a poly(A) or repetitive A-rich tail at the 3' end, and a duplication of 7–20 base pairs (bp) of target sequence flanking the insert. *Alu*-like sequences, processed mRNA pseudogenes and snRNA pseudogenes all show these features, and seem likely to have a common mode of origin⁵. The term 'retrotransposons' has been suggested for all such sequences⁶. Sequences of LINEs from primates and rodents show the same features, implying that the LINEs are also retrotransposons.

The rodent LINE, called MIF-1, had earlier been restriction-mapped by several groups², but its sequence has emerged by a

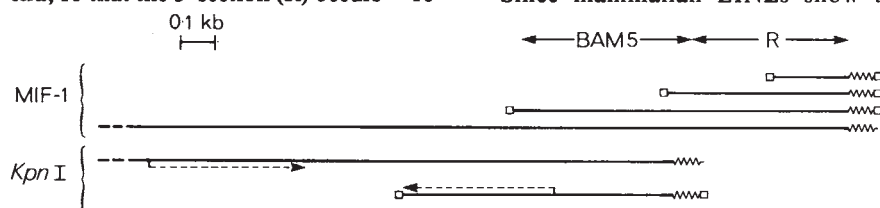
circuitous route. Rodent genomes contain several families of short *Alu*-like retrotransposons, and when two groups^{7,8} defined the 0.5-kilobase (kb) 'R' repeats, they seemed to constitute one more such family. Both R and the *Alu*-like sequences exist in ~10⁵ copies per genome and show the typical retroposon hallmarks. However, the 5' ends of the R sequences were of variable length⁷, and several extended⁹ into a repeated sequence previously named BAM5 (ref. 10). This is itself commonly part of an even larger repeated sequence, which has now turned out to be MIF-1 (refs 11,12; see the figure). Only fragments of the larger MIF-1 sequence have been published so far^{12–15}, but the deduced continuity is confirmed by homology with the human equivalent¹². Thus, MIF-BAM5-R is a single sequence of about 7 kb. But whereas the major part of it occurs ~3 × 10⁴ times in the mouse genome, most copies are truncated at the 5' end, so that the 3' section (R) occurs ~10⁵

times per genome^{11,15}. Such truncation is also seen, rarely, with *Alu*-like retrotransposons, and it may simply reflect a tendency for reverse transcription beginning at the 3' end to terminate prematurely.

The primate equivalent is called *KpnI*. It too is about 6–7 kb long, and large parts of several examples have now been sequenced^{16–19}. Again, all copies sequenced end in a variable A-rich tail typical of retrotransposons, and many copies seem to be 5'-truncated^{17,20}; the latter include one clone of repeated DNA isolated several years ago²¹ but only now identified, and one insert in the monkey α -satellite DNA^{16,17}. Longer inserts in both monkey and human α -satellite DNA^{20,22} are also *KpnI* sequences. One 5'-truncated specimen from α -satellite has been fully sequenced^{16,17}, along with the 14-bp target site duplication which flanks it. However, it is an atypical specimen, as shown in the figure, since it includes a separate part of the *KpnI* sequence, which was inverted and joined to the 3' segment in a stretch where 9 out of 12 bases were fortuitously identical. Rearrangements may also be frequent in MIF-1 sequences, since Gebhard and Zachau have reported a tandem duplication within one BAM5-R unit⁷, as well as two separate cases where a BAM5-R unit is inserted into a pre-existing MIF-1 (ref. 11).

Monkey *KpnI* shows 60–70 per cent homology with murine MIF-BAM5-R, although the R portion (the last 0.5 kb) is missing from *KpnI*¹². The divergence between monkey and mouse is slightly less than that of redundant DNA such as introns, so the LINE sequence has been maintained, albeit loosely, either by selection or by correction. This of course refers to the consensus sequence. That the thousands of copies throughout the genome do adhere to a species-specific consensus, even while the consensus evolves, is a striking example of concerted evolution. It has already been shown that LINEs evolve in concert within rodent species²³ and within primate species²⁰; now it appears that this has been going on at least since these two orders diverged (ref. 12 and references therein). It remains to be seen how much of this is due to gene conversion of copies *in situ*, and how much is due to turnover of copies by insertion and deletion at different locations. The mammalian LINEs show no evident homology with an otherwise similar LINE family in *Drosophila*, which also shows target-site duplication and heterogeneous 5'-truncation^{23,25}.

Since mammalian LINEs show the



Alignment of the 3'-terminal parts of rodent and primate LINEs. Zigzags indicate variable A-rich tails; boxes indicate duplications of the insertion site (not yet demonstrated for full-length LINEs). The *KpnI* shown includes a separate inverted fragment of *KpnI* within it (dotted arrows)^{16,17}.

hallmarks of reverse transcription, one would expect there to be corresponding RNAs from which the reverse transcripts could be derived. Both MIF-1 and *KpnI* are indeed transcribed^{10,17,27,28}, and have been obtained as cDNA clones^{17,19,29}, which have the same 3' end as the genomic copies¹⁹. Most of the transcripts are made by RNA polymerase II, and most come from only one strand, as expected from the retroposon structures; but most of the transcripts are not polyadenylated, and are restricted to the nucleus²⁸. These transcripts may well reflect considerable internal complexity in the LINES. For example, *KpnI* hybridizes to several discrete cytoplasmic RNAs which are shorter than the complete LINE²⁷, while MIF-1 contains homology to a mobile insertion sequence (the so-called 'IAP LTR') which carries a promoter³⁰. But it is not yet clear how the transcripts are organised, nor even whether they are directed by the LINES themselves, rather than by promoters in flanking sequences of some copies.

In fact, it remains possible that the LINE retroposons, although comprising the great majority of LINES, are processed pseudogenes derived from some unidentified parental copy or copies^{17,19}. Further understanding must wait until some full-length specimens have been completely sequenced and their transcripts accurately mapped. Until then, the identification of the majority as a single family of retroposons at least means that another few percent of the mammalian genome has been accounted for. □

John Rogers is at the MRC Laboratory of Molecular Biology, Cambridge CB2 2QH.

Parasitology

Malaria and babesiosis

from F.E.G. Cox

MALARIA AND BABESIOSIS, two of the most important parasitic diseases in the world, affecting man and cattle respectively, are caused by protozoa that live and divide within red blood cells. Eventually, destruction of the red blood cells causes anaemia and initiates a range of pathological processes frequently leading to death. This intraerythrocytic habit links the causative organisms, *Plasmodium spp.* and *Babesia spp.* and a recent conference* brought together those involved in the medical, veterinary and scientific aspects of malaria and babesiosis in the hope that knowledge gained in one field would complement that obtained in the other.

The keynote was control and the single most discussed topic was the interplay between protein biochemistry, hybridoma technology and gene cloning in the development of potential vaccines. In this field, mainly because of the lead taken by the World Health Organization, malaria is a long way ahead.

There are three candidate targets for a malaria vaccine, the infective stages injected by a mosquito (sporozoites), the late dividing stages in the blood (schizonts and their products, merozoites, that invade fresh cells) and the sexual stages or gametes. The sporozoites of all malaria parasites, including *P. falciparum* and *P. vivax* of humans, carry a surface antigen with a molecular weight of 42–58,000 with a single immunodominant region that is conserved and repetitive¹. In *P. knowlesi* in monkeys the 42,000 antigen is restricted to the outer pellicular membrane although the precursor molecules are associated with internal structures, the rhoptries micronemes². The antigen consists of a repeat of 12 amino acids and a synthetic peptide analogous to the native antigen induces an antibody response against the peptide³. Whether or not this peptide is actually protective remains to be seen and it is currently being tested in squirrel monkeys.

Monoclonal antibodies against sporozoites of the human parasite *P. malariae* also cross-react with *P. knowlesi* sporozoites⁴ supporting the view that the results of experiments with monkeys are directly comparable with those from man. An invaluable spin-off from this kind of work is the development of a test for the rapid and accurate identification of sporozoites in ground-up mosquitoes collected in the field⁵.

The asexual blood stages of malaria parasites produce a bewildering array of antigens but it is now becoming clear that those important in protection are produced late in the division cycle in the red blood cell and are expressed on the merozoites. In *P.*

knowlesi, an antigen with a molecular weight of 140,000 coats the surface of merozoites causing them to agglutinate and thus preventing the invasion of new cells⁶ *et al.* On the other hand, a merozoite surface antigen of the rodent malaria parasite, *P. yoelii*, does not seem to be involved with protection in such a simple way⁷ whereas an internal one associated with the rhoptries does⁸. Progress in the further identification of protective antigens is likely as a result of success in cloning *P. falciparum* antigens in *E. coli*. Of particular interest is the finding that one such antigen consists of 11 repeated amino acids but the similarity to the sporozoite antigen does not as yet go beyond this⁹.

The sexual stages of the malaria parasite have also received considerable attention because it is known that antibodies against these block transmission in the mosquito. Monoclonal antibodies against sexual stages reduce but do not totally prevent the infection of mosquitoes and do so only in combination but not separately¹⁰. Surface antigens with molecular weights of 48,000, 55,000 and 250–300,000¹¹ or 53,000, 59,000 and 255,000 have been identified¹⁰. Some have already been cloned and expressed in *E. coli*¹² but these experiments are still at a very early stage.

Against this background of progress at the molecular level in malaria, our understanding of the immunology of babesiosis looks very dated. Those working with babesiosis had a head start with a successful vaccine consisting of

*The Second International Conference on Malaria and Babesiosis was held in Annecy, France, on September 19–22 1983.

1. Zavala *et al.*, New York University School of Medicine, New York.
2. Fine *et al.*, New York University School of Medicine, New York.
3. Oysin *et al.*, New York University School of Medicine, New York.
4. Cochran *et al.*, New York University School of Medicine, New York.
5. Ramsay *et al.*, Naval Medical Research Institute, Bethesda.
6. Miller *et al.*, National Institutes of Health, Bethesda.
7. Freeman & Holder, Wellcome Research Laboratories, Beckenham, UK.
8. Oka *et al.*, Case Western Reserve University, Cleveland.
9. Coppel *et al.*, Walter and Eliza Hall Institute of Medical Research, Melbourne.
10. Carter *et al.*, National Institutes of Health, Bethesda.
11. Vermeulen *et al.*, University of Nijmegen.
12. Smits *et al.*, University of Nijmegen.
13. Kahl, Walter and Eliza Hall Institute of Medical Research, Melbourne.
14. Ristic, University of Illinois.
15. Kuttler *et al.*, Washington State University.
16. Timms *et al.*, Tick Fever Research Centre, Brisbane.
17. Montenegro-James *et al.*, Instituto de Investigaciones Veterinarias, Las Delicias, Venezuela.
18. Mazier *et al.*, Groupe Hospitalier Pitié-Salpêtrière, Paris.
19. Palmer *et al.*, University of Hawaii, Honolulu.
20. Jepsen *et al.*, Statens Serum Institut, Copenhagen.
21. Thomas, World Health Organization, Geneva.
22. Smith, University of Illinois.
23. Peters, London School of Hygiene and Tropical Medicine.
24. Fleckenstein, Walter Reed Army Institute of Research, Washington.
25. Cosgriff, Walter Reed Army Institute of Research, Washington.

1. Singer, M.F. *Cell* 28, 433 (1982).
2. Singer, M.F. *Int. Rev. Cytol.* 76, 67 (1982).
3. Jagadeeswaran, P. *et al.* *Cell* 26, 141 (1981).
4. Van Arsdell, S.W. *et al.* *Cell* 26, 11 (1981).
5. Sharp, P.A. *Nature* 301, 471 (1983).
6. Rogers, J.H. *Nature* 301, 460 (1983).
7. Gebhard, W. *et al.* *J. molec. Biol.* 157, 453 (1982).
8. Lueders, K.K. & Paterson, B.M. *Nucleic Acids Res.* 10, 7715 (1982).
9. Wilson, R. & Storb, U. *Nucleic Acids Res.* 11, 1803 (1983).
10. Fanning, T.G. *Nucleic Acids Res.* 10, 5003 (1982).
11. Gebhard, W. & Zachau, H.G., *J. molec. Biol.* (in the press).
12. Singer, M.F. *et al.* *Nucleic Acids Res.* (in the press); Bennet, K.W. & Hastie, N.D. *EMBO J.* (in the press).
13. Brown, S.D.M. & Plechaczky, M. *J. molec. Biol.* 165, 249 (1983).
14. Brown, S.D.M. *Gene* 23, 95 (1983).
15. Fanning, T.G. *Nucleic Acids Res.* 11, 5073 (1983).
16. Thayer, R.E. & Singer, M.F. *Molec. cell. Biol.* 3, 967 (1983).
17. Lerman, M.I., Thayer, R.E. & Singer, M.F. *Proc. natn. Acad. Sci. U.S.A.* 80, 3966 (1983).
18. Manueldis, L. *Nucleic Acids Res.* 10, 3211 (1982).
19. DiGiovanni, L., Haynes, S.R., Misra, R. & Jelinek, W.R. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
20. Grimaldi, G. & Singer, M.F. *Nucleic Acids Res.* 11, 321 (1983).
21. Deininger, P.L. *et al.* *J. molec. Biol.* 151, 17 (1981).
22. Potter, S.S. & Jones, R.S. *Nucleic Biol.* 150, 441 (1981).
23. Brown, S.D.M. & Dover, G. *J. molec. Biol.* 150, 441 (1981).
24. Rae, P.M.M. *Nucleic Acids Res.* 9, 4997 (1981).
25. Dawid, I.B. & Rebertus, M.L. *Nucleic Acids Res.* 9, 5011 (1981).
26. Rolha, H. & Glover, D.M. *Nucleic Acids Res.* 9, 5521 (1981).
27. Kole, L.B., Haynes, S.R. & Jelinek, W.R. *J. molec. Biol.* 165, 237 (1983).
28. Shafit-Zagardo, B. *et al.* *Nature* 304, 277 (1983).
29. Soriano, S., Meunier-Rotival, M. & Bernardi, G. *Proc. natn. Acad. Sci. U.S.A.* 80, 1816 (1983).
30. Brown, A. & Huang, R. *Proc. natn. Acad. Sci. U.S.A.* 79, 6123 (1982).

attenuated blood forms for which the factors determining virulence and transmissibility are only just being understood¹³. The development of *in vitro* culture techniques has made it possible to prepare a vaccine from soluble antigens released into the culture media. So far, three antigens in the molecular weight range 37–40,000 have been identified¹⁴ and these have been used as vaccines against homologous^{15,16} and heterologous¹⁷ challenge but with no more success than with the conventional vaccine.

Although immunology dominated the meeting significant progress was reported in other areas, such as *in vitro* cultivation. Of particular interest was the successful cultivation of the liver stages of *P. vivax* in hepatocytes¹⁸ and the reports of mass cultivation techniques for *P. falciparum*^{19,20}.

In epidemiology, the use of computer simulations of field situations for both malaria²¹ and babesiosis²² has reached very sophisticated levels and will improve our understanding of the complex nature of these diseases — a prerequisite for any

comprehensive control strategy. However, the main weapon of control, chemotherapy, is in a very poor state. With increasing resistance to antimalarials, including Mefloquine, which has not been widely released²³, clinicians are having to rely on old remedies such as quinine. Possible new antimalarials including Halofantrine²⁴ and Pyridinemethanol (WR180, 409)²⁵ show promise as do derivatives of the Chinese folk drug artemisinin such as artesunate which is many times more effective than quinine²³. Drug resistance has also been experienced in babesiosis¹⁵ but there are no new drugs on the way. One wonders what the situation in chemotherapy would have been like if a fraction of the effort devoted to immunology had been diverted in its direction. One thing is certain, though, malaria and babesiosis will still be with us when the third conference is held. □

F.E.G. Cox is Professor of Zoology at King's College London, Strand, London WC2R 2LS

Astronomy

Formation and fate of interacting binaries

from Virginia Trimble

BINARY stars as a subject of scientific inquiry are, in one sense, very old (ancient and boring, many of our colleagues in extragalactic astronomy would say), having been invented by John Michell¹ more than two hundred years ago when he pointed out that the number of close pairs of stars in the sky is so large that most of them must be physically connected and not just chance superpositions. In another sense, the subject of interacting binaries is quite young. The first international meeting devoted exclusively to them (International Astronomical Union Colloquium no. 6) took place in September 1969 in Elsinore, Denmark, and the most recent (a NATO Advanced Studies Institute) in August 1983 in Cambridge, UK. None of the official participants in the former was present at the latter, marking 14 years as the length of a generation of binary-star astronomers, at least in their guise as conference goers.

During this recent historical period, our physical understanding of interacting binaries has tended to proceed from the middle evolutionary phases outwards in both directions. The first systems to be put on a firm theoretical foundation were the Algols², which represent the stage when the more rapidly evolving component has nearly completed the process of transferring material to its companion. Next came the cataclysmic variables^{3,4} and X-ray binaries^{5,6}, in which the first star has exhausted its nuclear fuel and is accreting material transferred back from the

secondary. Then the phases just before and during rapid transfer, termed RS CVn^{7,8} and Serpentiid⁹ systems, were identified; and plausible models were calculated¹⁰ for very close pairs of main-sequence stars sharing a common envelope (W Ursa Majoris variables).

Now we are finally beginning to sneak up on the earliest and latest phases. The result is not always increased clarity. We used, for instance, to believe that the fission of a single rotating gas mass would make close pairs with nearly identical components, while separate condensation would make wide pairs with the primary much more massive than the secondary^{11–13}. Some recent models work this way¹⁴. Others do not. Gingold and Monaghan¹⁵ have found that fission of a differentially rotating cloud can produce close pairs with mass ratios very different from one. And P. Artymowicz (University of Warsaw) reported calculations of separate accretion (on to cores of 0.2 solar masses resulting from hierarchical fragmentation) whose most likely product is a close pair with total mass 1–3 solar masses and mass ratio near one.

At the other end of the evolutionary track, R.E. Nather (University of Texas, Austin) discussed the close pairs of white dwarfs that seem to be descendants of cataclysmic systems. There are now three (AM CVn, GP Com and PG1346+082), raising them to the status of a well known class of astronomical object. Systems like them are strong candidates for the immediate

progenitors of type I supernovae^{16–18}. All three were found by accident. Thus close binary white dwarfs could be very common, and one of the questions left unanswered at the end of the meeting was how to search efficiently for the ones in which the stars are not interacting.

The progenitors of cataclysmic variables are seen in rather larger numbers than the products. H. Bond (Louisiana State University) showed data for about a dozen promising systems. That is, they are close, but non-interacting, pairs with one main-sequence component and one that is either ionizing a planetary nebula or identifiable as a white dwarf or 0 subdwarf. 'Close' means both that the stars must have spiralled together since the white dwarf was a red giant and that loss of angular momentum by gravitational radiation and/or winds will cause them to interact within another Hubble time. These systems appear to be the short-period tail of a more extensive class. Bond also listed a somewhat larger number of WD+MS systems with periods greater than 1.5 days. These are unlikely to come into contact in the age of the Universe, but are too close to have evolved without extensive loss of angular momentum.

Most of the necessary angular momentum loss appears to have occurred during a phase when the stars share a common envelope¹⁹. R. Taam (University of Illinois) presented results of two-dimensional calculations of binary systems enveloped by the expansion of a red giant. Mass outflow occurred at rates as large as 10^{26} g s^{-1} , mostly in the equatorial plane. Complete expulsion of the envelope occurred most readily in wide systems (because the red giant was more extended and its envelope less tightly bound when the process started) and took only a few years. No wonder we don't catch many systems in the act.

We still cannot say that we understand the evolution of binary systems from symmetry breaking in the early Universe to the boiling off of black holes by Hawking radiation. But progress is being made. □

Virginia Trimble is Visiting Professor of Astronomy at the University of Maryland, College Park, Maryland 20742, and Professor of Physics at the University of California, Irvine.

1. Michell, J. *Phil. Trans. R. Soc.*, p.234 ff (1767).
2. Paczyński, B. *A. Rev. Astr. Astrophys.* **9**, 183 (1971).
3. Robinson, E.L. *A. Rev. Astr. Astrophys.* **14**, 119 (1976).
4. Warner, B. *IAU Symp.* **73**, 85 (1976).
5. Blumenthal, G.R. & Tucker, W.H. *A. Rev. Astr. Astrophys.* **12**, 23 (1974).
6. van den Heuvel, E.P.J. *IAU Symp.* **73**, 76 (1976).
7. Webbink, R. *IAU Symp.* **88**, 395 (1980).
8. Hall, D.S. *IAU Colloq.* **29**, 287 (1976).
9. Plavec, M.J. *IAU Colloq.* **69**, 172 (1982).
10. Shu, F. *IAU Symp.* **88**, 477 (1980).
11. Opik, E. *Publ. Obs. Astr. Univ. Tartu* **25**, no.6 (1924).
12. Bleksley, A.E.H. *Nature* **133**, 613 (1934).
13. Blaauw, A. & van Albada, T.S. *IAU Symp.* **30**, 215 (1967).
14. Lucy, L.B. & Ricco, E. *Astr. J.* **84**, 401 (1979).
15. Gingold, R.A. & Monaghan, J.T. *Mon. not. R. astr. Soc.* **204**, 715 (1983).
16. Paczyński, B. Preprint (1983).
17. Webbink, R. Preprint (1983).
18. Iben, I. & Tutukov, A.V. Preprint (1983).
19. Paczyński, B. *IAU Symp.* **73**, 65 (1976).

Psychology

Widening vistas for eye movement research

from Peter Coles

UNTIL recently the study of human eye movements was strictly for the psychology laboratory; recording required the fitting of uncomfortable contact lenses or the use of head restraints, machines were unreliable, and analysis had to be carried out by an exceedingly tedious manual procedure. Now, unobtrusive computerized machines are becoming available and are opening up eye movement research to workers from a whole new range of disciplines. Some of the new developments — in areas as diverse as neurophysiology and quality control on the production line — were reported at a recent European conference.*

The new machines make it possible to study eye movements in children, and even infants, and provide vital new information for the study of both normal and abnormal visual development. Of particular interest is the possibility that abnormal oculomotor behaviour may either reflect, or be a contributing factor, in reading disorders. One study suggested that children often show inaccurate convergence of the two eyes so that the two foveas are not aimed precisely at the same target (J.F. Stein and S. Fowler, University of Oxford). The authors developed the hypothesis of a 'leading eye' which acts to suppress the conflicting image of the other eye and enables more reliable associations between ocular motor and retinal signals to improve the accuracy of binocular eye movements, in much the same way as Walls suggested in his theory of ocular dominance in 1950. A binocular test showed that by age five years, over 50 per cent of children sampled had developed a stable 'leading eye', whilst the figure was over 90 per cent by age eleven years. It was found that those children who had not developed a stable leading eye had reading ages as much as eight months behind those who had developed a stable leading eye. These results were put to use by investigating effects on poor readers' performance of wearing spectacles which occluded vision from one eye, thus simulating the function of a 'leading eye'. Children with visual dyslexia (and an unstable leading eye) were divided into two groups, one wearing the occluding spectacles for all close work, the other not. After six months, the dyslexic sample who had worn an occluding lens had improved their average reading age by twelve months, whilst the control sample had improved their average reading age by six months. About 50 per cent of the experi-

mental group had also developed a stable leading eye.

Knowledge about the sequence and localization of ocular fixations as opposed to the movements themselves, can help understand, and perhaps improve, efficiency of routine, but critical visual inspection tasks, such as scanning chest X-ray images for tumours or signs of disease. Up to 30 per cent of small tumours recorded on chest images are not reported, even by trained radiologists. Performance cannot be improved simply by presenting sections of the image serially, possibly because this prevents the comparison of potential lesions or 'nodules' with areas judged to be healthy, as occurs in free viewing (D.P. Carmody, St. Peter's College, Jersey City). The efficiency with which radiological examinations are carried out varies considerably during the day. Ocular fixations are of significantly longer duration in morning sessions than after lunch or at the end of the working day (A. Gale, Queen's Medical Centre, Nottingham).

Recent developments in the study of eye movements themselves were mostly concerned with the rapid 'voluntary' saccadic eye movements. The saccadic system is traditionally characterized as being

'ballistic' in so far as the movement is pre-programmed and cannot be altered in mid-flight. This preprogramming is generally thought to occur in the necessary intersaccadic intervals which seem to be characterized by a minimum refractory period before another saccade can be initiated, with vision being 'suppressed' during the saccade itself. Modern recording techniques make it possible to change the stimulus display as a function of ongoing eye movement and research using these 'contingent' paradigms has led to modifications of most of the 'traditional' characterizations of saccades.

It has emerged that saccade peak velocities are not simply determined by target eccentricity, as had been assumed, but may be slowed in mid-flight (T.J. Crawford, University of Durham). Further evidence was offered for the idea of a speed/accuracy trade-off in the saccadic system, with the implication that accurate saccades may involve a different mode of use of sensorimotor pathways (Z. Kapoula, University of Paris). A very useful contribution was made to our understanding of changes in oculomotor functioning in early infancy, a subject-group which has proved particularly elusive as long as recording devices required attachments to the eye or orbital areas. Whilst infants under five-months of age showed saccades having velocity parameters similar to adults, a subset of infants also showed 'back-to-back' saccades with negligible inter-saccadic intervals, probably never occurring in normal adults (L. Hainline, J. Turkel & I.A. Abramov, Brooklyn College of

Astronomy

1983 TB and the Geminids

from David W. Hughes

A fast-moving Apollo-type asteroid, 1983 TB, has recently been discovered with the infrared astronomy satellite IRAS. The asteroid was then tracked by observers at the Palomar and Lowell observatories and the orbit calculated by C.M. Bardwell at the Center for Astrophysics, Cambridge, Massachusetts (see IAU circular no.3879). Two surprises were in store. 1983 TB has the smallest perihelion distance of any known asteroid and, even more exciting, the orbit is nearly coincident with that of the Gemini meteor stream. The parameters are given below.

A recent paper on the orbital evolution of the Geminids investigated the possibility that an unknown minor planet could have been influencing the stream (*Mon. Not. R. Astr. Soc.* **199**, 313; 1982). The only asteroid that was known to come anywhere

near was 132 Aethra and it was estimated that this only had a minimal effect. Now the new observations show there is a minor planet actually in amongst the meteoroids.

Others have often speculated that comets decay to produce both a stream of meteoroid dust particles and a dead cometary nucleus — this being indistinguishable from an asteroid. One problem was that most known asteroids are considerably larger than cometary nuclei. But still the orbits of Encke's comet, the daytime Taurid meteor stream and asteroid Hephaistos were reasonably similar. Those of the Geminids and 1983 TB are much much closer. □

David W. Hughes is a Lecturer in Astronomy and Physics at the Department of Physics, University of Sheffield, Sheffield S3 7RH.

	Ω (Deg)	ω (Deg)	i (Deg)	e	q (AU)	a (AU)	P (Years)
Geminids	260.3	324.8	23.6	0.896	0.140	1.35	1.57
1983 TB	261.83	324.59	22.77	0.8944	0.1376	1.3027	1.49

*Second European Conference on Eye Movements, Queen's Medical Centre, Nottingham, 19-23 September 1983. Organized by the Biological Engineering Society in association with the Applied Vision Association.

CUNY, New York).

Developments in the neurophysiology of eye movements still relied largely on evidence from animal studies or human pathological studies, with the consequent problems of generalisation. Experiments using monkeys provided evidence for 'express saccades', having latencies as short as 75 msec and also contributed to our knowledge about neural mechanisms involved in convergence and accommodation. The possibility of eye movement

recording outside the laboratory has permitted various industrial inspection and quality-control tasks to be studied. Those reported included the inspection of apples for blemishes and the inspection of roller bearings for faults, but the results do not yet seem to have direct consequences for improving performance. □

Peter Coles is at Wolfson College, University of Oxford, Linton Road, Oxford.

Earth sciences

Geomagnetic impulses and deep mantle conductivity

from K.A. Whaler

IT now seems certain that around 1969 a spectacular change took place in the geomagnetic field. The change was almost synchronous over the whole of the Earth's surface, took place in less than two years, and is now known to have consisted of a 'jerk': a step change in secular acceleration of the magnetic field¹⁻³ that has its origin inside the Earth^{4,5}. A new paper from G.E. Backus⁶ investigates the propagation of magnetic impulses — of which the jerk is one example — through the electrically conducting mantle and provides a rigorous mathematical framework in which to investigate constraints on the mantle's electrical conductivity profile provided by jerk data. The results show once again how 'geophysical intuition' can lead us astray — observations of rapid changes in the geomagnetic field originating in the core, which one would tend to associate with weak electrical screening by the mantle, do not necessarily imply low values of mantle conductivity.

Backus regards the mantle as a linear filter which acts on impulses injected from the core, and whose output is observed at the Earth's surface. The filtering is different for each harmonic component of the signal; the most useful indicators are the zero-frequency delay time and smoothing time, both in principle wavelength dependent. The zero-frequency delay time is the time a given harmonic degree impulse takes to propagate through the mantle, and the smoothing time is its spread by the time it reaches the Earth's surface; Backus sets bounds for the ratio of the two timescales. Both characteristic times are functionals of the conductivity profile of the mantle — the delay time is linearly related, and the smoothing time quadratically related, to the conductivity — so estimates of them could provide the data for an inversion scheme (although there are no standard non-linear procedures).

The magnetic observatory records from Western Europe provide the clearest indication of the jerk². Data from a single area

of the Earth's surface do not, however, provide any spatial resolution so do not permit wavelength content to be analysed. Backus therefore interprets the European data in terms of a fictitious 'mixed' filter, which combines all the important, low harmonic degree filters. The output from such a 'mixed' filter can have a smoothing time shorter than that of any of the constituent filters, and explains why observed rapid magnetic changes do not require low mantle conductivity values.

More detailed studies using globally distributed data are needed to determine the parameters of individual mantle filters, which are more closely related to mantle conductivity values. Worldwide, a picture similar to that seen in Western Europe emerges⁵, suggesting that the delay and smoothing times are similar for each mantle filter. The main effect of the conducting mantle is, then, to impose a uniform delay on the core field observed at the Earth's surface; this means that the spherical harmonic coefficients of the field we see today are those of the field the geodynamo generated a time τ_1 ago, where τ_1 is the (common) delay time of the mantle filters.

One crucial question remains; what is the starting time of the impulse at the core? The answer cannot come from magnetic data alone and, without it, the delay time of the mantle filters cannot be determined from the emergence time of the signal. There is some evidence, however, of a correlation between changes in the magnetic field and changes in the length of day⁷⁻⁹, assumed to reflect changes in the balance of angular momentum between the core and mantle, such that the total for the Earth is conserved. The date of the length of day change which correlates with the 1969 magnetic jerk could be taken as the starting time of the impulse at the core. Backus identified an 'elbow' in the length of day curve¹⁰ in 1956 as the date of the impulse — a highly controversial decision and one on which inferences concerning the conductivity of the mantle depend

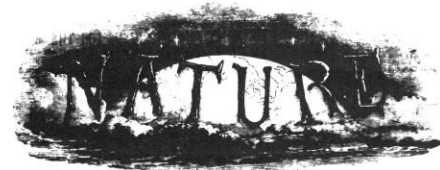
strongly.

There are two main problems with Backus' choice. First, previous calculations based on long time spans suggest that length of day changes lead magnetic changes by about five years^{7,8}, not thirteen years, while others find length of day lagging magnetic changes by a similar amount⁹. Second, not every published length of day curve shows an elbow in 1956. Many geophysicists might perhaps prefer to regard Backus' calculations as merely illustrative numerical examples!

Despite the very great importance of this problem Backus' theory gives investigations of core impulses a firm quantitative basis. Clearly, mantle conductivities will be difficult to extract, but with the current state of knowledge even quite general results would represent a considerable step forward. The jerk is one of the most exciting recent geomagnetic events, and Backus' paper demonstrates how to maximise the possible information from it. □

K.A. Whaler is in the Department of Earth Sciences, The University, Leeds LS2 9JT.

1. Courtillot, V., Ducruix, J. & Le Mouél, J.-L. *C. r. heb. Séanc. Acad. Sci., Paris* **287**, 1095 (1978).
2. Achache, J., Courtillot, V. & Le Mouél, J.-L. *Phys. Earth Planet. Interiors* **23**, 72 (1980).
3. Ducruix, J., Courtillot, V. & Le Mouél, J.-L. *Geophys. J. R. astr. Soc.* **61**, 73 (1980).
4. Malin, S.R.C. & Hodder, B.M. *Nature* **296**, 726 (1982).
5. Malin, S.R.C., Hodder, B.M. & Barraclough, D.R., *Ebro Observatory 75th Anniversary Volume* (in the press).
6. Backus, G.E. *Geophys. J. R. astr. Soc.* **74**, 713 (1983).
7. Vestine, E.H. & Kahle, A.B. *Geophys. J. R. astr. Soc.* **15**, 29 (1968).
8. Kahle, A.B., Ball, R.H. & Cain, J.C. *Nature* **223**, 165 (1969).
9. Le Mouél, J.-L., Madden, T.R., Ducruix, J. & Courtillot, V. *Nature* **290**, 763 (1981).
10. Morrison, L.V. *Geophys. J. R. astr. Soc.* **58**, 349 (1979).



100 years ago

M. Raphael Perulta writing to *La Nature* under date Manila, September 14, states that the detonations of the Java eruption of August 27 were distinctly heard throughout the Philippine Islands; so distinctly were the sounds heard that gunboats were sent out under the impression that a fight was going on at Java, or that a ship in distress was firing for help.

On October 26 at about 7 p.m. a splendid meteor was seen in the district of Hernö and, Sweden. A traveller on the road to Ragunda states that he suddenly saw the night lit up as in broad daylight, which was caused by a large meteor appearing with a blinding white lustre in the zenith and travelling very rapidly down to the horizon. When half way, as it appeared to the observer, between zenith and the earth it suddenly burst, throwing a quantity of sparks in every direction.

From *Nature* **29**, 44 (1883).

Supernovae and life

SIR — We call attention to synchrotron radiation emanating from supernovae as a source of circularly polarized light capable of inducing asymmetric photochemical reactions of prebiotic molecules. This light is predominantly polarized in the plane of motion of fast electrons orbiting around supernovae and when viewed off-axis the polarization is elliptical or circular. Chiral molecules in the interstellar medium or on planetary surfaces would undergo preferential photosynthesis or photolysis when irradiated by such a source. Domains on opposite sides of the plane of predominant polarization would be exposed to light of opposite helicity so that overall symmetry would be preserved. But in a single locale, one enantiomer would predominate.

Molecular chirality can be approached either as an expression of random fluctuation or as the result of a specific mechanism. Mann and Primakoff point out that statistical fluctuation as the basis of chiral dominance is highly unlikely because it would require the assumption of a very small number of terrestrial polymerization sites¹. The physical processes that have been invoked include the interaction of a racemic mixture with electrons of specific helicity emitted by a nuclear β -decay process, and the apparently more effective interaction of circularly polarized light (CPL) with molecular orbitals. Photosynthetic and photodestructive reactions with CPL have been described in a number of reviews, and a 2.50% optical enrichment of a racemic mixture of the amino acid leucine has been achieved by photolysis with 212.8-nm CPL².

The sources of CPL previously considered include sunlight reflected at the Earth's surface, and by the Earth's magnetic field². Alignment of interstellar grains has also been identified as a source of linear polarization, and a phase shift can lead to circular polarization³. These mechanisms require two independent processes: scattering or preferential absorption, followed by optical rotation in a birefringent medium of well-defined optical axis and proper optical thickness.

Synchrotron radiation is also an important source of CPL. The history and the properties of synchrotron radiation have been recently reviewed by Winick and Doniach⁴. Liénard first discussed the concept of radiation from a charged particle moving circularly. Sokolov and Schwinger provided theoretical treatments and Pollock and his colleagues made some of the early experimental studies⁴.

Shklovskii suggested that the continuous light from the Crab Nebula arises from the synchrotron process, Dombrovskii demonstrated that the light is polarized, and detailed studies in the optical range were undertaken by Oort and Walraven and by Baade⁵. These investigations showed that electrons are continuously

ejected by the supernova remnant and that the synchrotron radiation persists after the initial explosion. Although stable polarization is a characteristic of some sources, in others there are fluctuations and multiple components.

A number of lines of astrophysical evidence link supernovae to the formation of planetary systems⁶. The intensity, the spectrum (extending into the X-ray region), and the high degree of polarization, including circular polarization, of supernova synchrotron radiation, indicate that this light is capable of processing polyatomic molecules in nearby interstellar matter, leading to asymmetric photosynthesis or photolysis in out-of-plane domains. The circular polarization is a direct effect that does not require fortuitous alignment of consecutive optically active media.

For life to arise from chiral building-block molecules modified by synchrotron light, it would be necessary for these either to survive intact during the heating process of planet formation or to be ferried on incoming grains as the planet periodically traverses clouds of interstellar material or to be assembled on tepid planetary surfaces illuminated by residual supernova synchrotron radiation.

E. RUBENSTEIN
W.A. BONNER
H.P. NOYES
G.S. BROWN

Stanford University
Stanford,
California 94305, USA

1. Mann, A.K. & Primakoff, H. *Origins of Life* (in the press).
2. Flores, J.J., Bonner, W.A., & Massey, G.A. *J. Am. Chem. Soc.* **99**, 3622-3624 (1977).
3. Savage, B.D. & Mathis, J.S. *A. Rev. Astron. Astrophys.* **17**, 73-111, (1979).
4. Winick, H. & Doniach, S. in *Synchrotron Radiation Research* (eds Winick, H. & Doniach, S.) 4-6; 11-25 (Plenum, New York, 1980).
5. Baade, W. *Bull. Astron. Instit. Netherlands* **12**, 312 (1956).
6. Herbst, W. & Assousa, G.E. *Scient. Am.* **241**, 138-145 (1979).

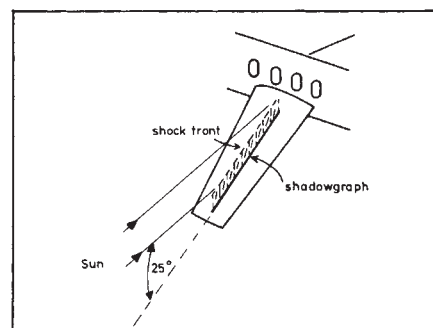
In-flight movies

SIR — Shadowgraphs provide a simple method of studying aerodynamic shock waves, but I was unaware until recently that striking shadowgraph patterns may sometimes be seen on the wings of civil aircraft. I understand that this phenomenon is wellknown within the aerodynamic fraternity but my observations may interest others who, in the spirit of Jearl Walker¹, have a taste for science outside the laboratory.

At cruising altitude, the flow over the leading edges of the wings of conventional airliners is accelerated to trans-sonic speeds. This may give rise to shock fronts which stand perpendicular to the wing-surface along its span, and extend upwards for 1 or 2 metres. Light traversing the shock at glancing incidence is refracted towards

the trailing edge of the wing and may produce a shadowgraph. Lamplough² has demonstrated this effect in flight using collimated light from an artificial source directed parallel to the wing-surface.

I recently observed similar shadowgraphs cast by natural sunlight while flying at an altitude of about 10⁴m in a Boeing 727. The shadow-band extended in an almost straight line along most of the width, as sketched in the diagram, and I estimated its width (by comparison with rivets) to be roughly 1-2 cm. A narrower bright band, akin to a caustic, lay along the rear boundary of the shadow. The pattern was of moderate to low contrast seen against the sunlit wing surface, and it remained visible for more than an hour showing that the angle of illumination was not critical. The Sun was at an elevation of some 20°-30° above the wing-tip. The pattern occasionally jumped irregularly fore



and aft during periods of mild turbulence and multiple bands were sometimes seen. As the aircraft descended, the pattern shifted steadily towards the leading edge of the wing, fading and narrowing until it disappeared. This behaviour is consistent with a weakening shock due to reduced air speed.

In typical conditions, the Mach number of the flow might change from 1.3 to 0.9 across the shock, producing a jump in density of about 0.1 kg m⁻³ and a corresponding increase in refractive index of about 2 × 10⁻⁵. Light traversing the shock at glancing incidence will be refracted through an angle of roughly 0.4° which is comparable with the angular size of the Sun; appreciable blurring of the shadowgraph may thus be expected. The width of the observed pattern showed that in my case the shock must have extended about 1 m above the wing.

High-flying observers wishing to see these effects and unacquainted with the characteristic appearance of shadowgraphs are recommended to consult photographs such as Figs 223 and 224 of ref 3. I am indebted to Dr P.J. Bryanston-Cross for illuminating discussions about trans-sonic flow near aerofoils.

A. HEWISH

Cavendish Laboratory,
Madingley Road,
Cambridge CB3 0HE, UK.

1. Walker, J. *The Flying Circus of Physics* (Wiley, New York, 1975).
2. Lamplough, A.E. *Aircr. Engng* **23**, 94 (1951).
3. Van Dyke, M. *An Album of Fluid Motion* (Parabolic Press, Stanford, 1982).

Popularizing science

It is more important today than ever, Isaac Asimov argues, that discoveries in science reach as broad an audience as possible.

IN 1686, a middle-class Frenchman, Bernard le Bovier de Fontenelle, was a failure. He had entered his father's profession of the law and it hadn't worked out. He decided to be a writer and had tried his hand at poetry, opera and drama, and had done poorly in each case.

But now, at the age of 29, he published a book entitled *Conversations on the Plurality of Worlds* and with that he hit the jackpot. It was an introduction, for the interested and intelligent layman, to the new astronomy of the telescope; a careful consideration of each of the planets from Mercury to Saturn, with speculations as to the kind of life that might be found upon them. The book was devoured by the upper and middle classes (who alone were literate) and went through many editions. Fontenelle was perhaps the first person to make a reputation in science on the basis of his popular science writing alone.

He was elected to the French Academy in 1691 and became perpetual secretary of the French Academy of Sciences in 1697. He wrote annual summaries of its activities, and composed obituaries for famous scientists as they died. He loved society and found himself welcome everywhere. He lived in good health and kept his faculties into advanced old age, dying at length one month before his 100th birthday.

There have been science popularizers ever since, including some important scientists — from Davy and Faraday, through Tyndall, Jeans and Eddington, to our contemporaries, Sagan and Gould.

But why bother? Few popularizers become wealthy, and there is certainly no great demand for it by the general public? Fontenelle wrote in the 'Age of Reason' and it was then chic for the aristocracy to allow themselves to be instructed in the new science. But now?

Well, science popularization today is essential, far more so than at any time in the past. Where, in Fontenelle's time, science popularization was entertainment for the otherwise-bored, it is now life and death for the world, for four reasons:

1) Science, together with its practical sister, technology, have taken over the world, both for better and for worse. Advances are coming at an ever-accelerating pace and the world is changing with breath-taking speed. If we consider only the last 40 years, we have television, the jetplane, the rocket, the transistor, the computer, each of which has completely altered the world we live in. Lying immediately ahead is the robotization of industry, the computerization of the home, and the expansion into space.

All these things, used wisely, can bring about (possibly) a much happier world than the one we live in now, or have ever lived in; a world more secure, more comfortable, more amusing. All these things, used unwisely, can destroy civilization.

We must not view science and technology as either an inevitable saviour, or an inevitable destroyer. It can be either, and the choice is ours. A general public, utterly ignorant of science, led by rulers scarcely better informed, cannot be expected to make intelligent choices in this matter. The alternatives of salvation and destruction depend, in that case, upon the blind gropings of ignorance.

This is not to say that we must build a world of scientists. That would be beyond the abilities of popularizers. But at least let the public make up an intelligent and informed audience. Football games are watched by millions who cannot play the game themselves, or even direct one successfully — but who can at least understand enough to applaud and to groan at the proper places.

2) Science is no longer the province of the well-to-do amateur who can finance his explorations out of his own pocket. The days of Henry Cavendish, Lord Rosse, and Percival Lowell are gone. Science is expensive and is growing more so. Continued advance is, to a large extent, dependent upon the support of large industries and of governments. In either case, it comes out of the public purse in the last analysis; and whether that purse will open or not, depends upon public willingness to endure the expense.

A public that does not understand how science works can all too easily fall prey to those ignoramuses like Senator Proxmire who make fun of what they do not understand, or to the sloganeers who proclaim scientists to be the mercenary warriors of today, and the tools of the military. The difference to the public between a scientist and a magician is the difference between understanding and not understanding and that is also the difference between respect and admiration on the one side, and hate and fear on the other.

Without an informed public, scientists will not only be no longer supported financially, they will be actively persecuted. Who, then, is to supply the understanding if not the science popularizers?

3) Scientists do not form a closed caste. They do not inherit their calling. New scientists must be recruited from outside, especially since the number of scientists, from decade to decade, is increasing at a more rapid rate than the number of people

generally is. How are they to be recruited?

Some youngsters are drawn to science willy-nilly by an inner compulsion, and cannot be kept out of it, but surely the numbers of these scientists-despite-themselves simply will not be great enough. There must be those who are attracted if some stimulus is applied, but perhaps not otherwise. An effective piece of science popularization is surely one way of rousing a spark of interest in a youngster; a spark which may eventually burst into flame.

I daresay there is not a science popularizer in the world who has not received a gratifying number of letters from young readers who explain that they are majoring in physics (chemistry, biology, mathematics, geology, astronomy) and that the original push came from a book that the popularizer had written.

4) There have always been, throughout history, class-differences in society. Even when a particular society labours to abolish those differences and to proclaim all individuals equal before the law, natural differences will arise. Some differences are physical. A keen eye, razor-sharp reflexes, a marvellous voice, a talent for writing or for acting, simply can't be handed out to everyone.

And some differences are imposed by circumstance. Of two people of equal potentialities, one may receive a better education than the other, or one may gain more practical experience than the other. And although the race may not always be to the swift, that would be the wise direction in which to place your bets.

As the world progresses and becomes more permeated with science and technology, as it becomes more 'high-tech', the type of education that will become ever more necessary for individual advance, will be that in science. We are, in the future, running the risk of creating a world of science-educated 'haves', whose children will automatically have the opportunity of such education; and science-uneducated 'have-nots', whose children are not likely to have the opportunity. If one believes, as I do, that a strongly stratified society has the seeds of instability and destruction in it, then we should labour toward making science-education as broadly-based as possible, so that as few people as possible are fated from birth to be part of an uneducated under-class.

I would like to conclude now by saying that it was a consideration of all these compelling reasons that led me some quarter-century ago to abandon my academic duties to become a full-time science writer, but I cannot. The truth is that I enjoy writing so much more than I enjoy anything else, that I would have made the switch if there were no other reason for it at all. □

Isaac Asimov's most recent book, The Roving Mind (Prometheus), was published this year and will be reviewed in Nature.

Fluid sinology

Johnathan Spence

Science and Civilisation in China: Vol.5, Part 5. Spagyric Discovery and Invention: Physiological Alchemy.

By Joseph Needham and Lu Gwei-Djen.

Cambridge University Press: 1983. Pp.572. £49.50, \$99.

THERE can be few things harder to study than the alchemical traditions of a particular society. Couched in evasive and ambiguous language, dealing by definition with complex transformations of matter and of mind, roving from precise laboratory experimentation to the wildest flights of cosmic fantasy, conducted often by marginal figures in the terms of conventional social mores, alchemical writings are at once mines of information and minefields for the unwary. Though thorough sifting of European alchemical materials for their scientific content was undertaken decades ago, Chinese alchemical writing proved — not surprisingly — all the more obdurate. Now with the appearance of the fifth part of volume 5 of *Science and Civilisation in China* Joseph Needham and Lu Gwei-Djen bring to an end their lengthy preliminary exploration of Chinese alchemical writing over the last two thousand years.

The previous three parts of volume 5 (on which Professor Nathan Sivin was a key collaborator) dealt with the tradition of what the Chinese called *wai-tan* experimentation, which Needham translates as "inorganic laboratory alchemy", in other words the making of elixirs of various types by the transformation of minerals (especially mercury), or the transformation of 'base metals' into silver or gold. In those remarkable volumes Needham explored the relationship between this *wai-tan* tradition and the development of experimental science in China. Now in the present volume he gives his attention to the inner or *nei-tan* branch of the Chinese tradition, which he translates generally as "physiological alchemy". Needham defines this as being the making of macrobiogens from living body juices, a process to which he gives the neologism "enchymoma". The purpose of making the enchymoma is to restore and preserve youth in the most fundamental of ways, by regenerating the tissue, returning to the pure state of infancy before decay. Needham refers to this deep "returning" with another neologism as being "anablastemic" — thus the phrase "anablastemic enchymoma" appears frequently in his key translations of Chinese passages as the topic of analysis by the Taoist practitioners whom, he believes, brought *nei-tan* alchemy to its greatest perfection.

Although some of the texts which Needham himself cites seem to incorporate in their descriptions of *nei-tan* such practices as the use of certain chemical elixirs, the development of health through solar and lunar rays, or esoteric diets served from unusual plants, Needham does not follow them. Instead he sees the pure *nei-tan* tradition as incorporating seven specific techniques, alone or in combination: bodily hygiene; respiration; fluid circulation; gymnastics; the preserving of



A *nei-tan* adept

internal secretions (especially saliva); certain sexual activities; and the state of trance. In the important and fascinating chapter four of the volume he explores these seven areas, offering an astounding range of textual examples, and examining them with thoroughness and — at times — with disarming wit. Thus he gravely declares that he feels he must let the "Chinese call a spade a spade, even if a jade one" (p.185) in lighthearted reference to the Chinese penchant for referring to the male sexual organ as the "jade stalk" ("jade flute" is often preferred in Chinese fiction and poetry); and he refers to the Taoist practice of trying to channel the male semen to the brain during intercourse with the wondrous phrase; *coitus thesaurus*.

Needham leads us into his analysis by way of a protracted examination of Jung's use of Chinese examples in the development of the theory of archetypes, a use that Needham rejects as ill-advised and

fundamentally inaccurate; and he accompanies it by reference to the traditions of European alchemy and to the Indian traditions of yoga, including the Hathayoga practices. In a final coda he discusses, as component to the *nei-tan* tradition, the uses of human urine or the placenta in Chinese medicine, which he interprets to show that the Chinese anticipated the Western development of urinary steroids and protein hormones.

The response of a historian without any extensive scientific training (such as myself) to this mass of data and argument is partly one of delight and awe — delight at having so many hitherto unknown texts brought to his attention, accompanied by the key terms in Chinese to help one check the accuracy of Needham's fluent but bold translations; awe at the hundreds of speculations Needham induces concerning the role of science within Chinese society, and those who practiced it. In part, however, the response is also one of scepticism and reserved judgement, as is proper with a work of this range and didactic purpose.

The reader has to ask himself, again and again, whether Needham is not over-translating, or at least giving the appearance of greater scientific specificity to the Chinese text than the Chinese text itself justifies. This argument will only be settled after decades of discussions between those qualified in both science and sinology, still a tiny number, but even in this present volume, and as an outsider, one can point to the fact that in discussing the beginnings of the *nei-tan* tradition Needham himself mentions that the text is a sixth century Buddhist one and the term appears as a "literary trope" (p.140), and at many other times he refers to the problems that arise from the cloudings of the *nei-tan* tradition by "symbolic language" or the use of "parable and metaphor" (i.e., pp.211 and 219). Brilliantly ingenious though Needham is in explaining how certain types of "scientific" illustrations that one could have thought lay in the *wai-tan* realm are in fact *nei-tan*, one can never be sure that much of the language may not be employed in a literary and figurative way.

Needham suggests that from the T'ang dynasty onward "physiological alchemy" developed apace because it was so much "cleaner" than "...messing about' with minerals, herbs and metals". (p.209). The judgment is tongue-in-cheek in tone, though seriously meant; one might add that just as *nei-tan* was a swerving away from *wai-tan* practices, so might a later generation have had excellent reasons for rejecting the *nei-tan* as well. □

Johnathan Spence is Chairman of the Department of History at Yale University. His book *The Memory Palace of Matteo Ricci (Viking)*, describing a Jesuit mission to China in 1783, is to be published shortly.

Wagons and chariots

A.F. Harding

The Earliest Wheeled Transport: From the Atlantic Coast to the Caspian Sea.
By Stuart Piggott. *Thames and Hudson:*
1983. Pp.272. £20, \$34.95.

THE WHEEL is one of those technological inventions which, in the banal and illogical catch-phrase, 'changed the course of history'. What would Old World civilization have been like without it? What would the New World have been like with it? Wheels are among those devices, such as pottery or writing, that are so basic one cannot easily imagine life without them. In view of this, it is surprising that no modern monograph has hitherto been devoted exclusively to the early history of wheeled transport, and a long-standing gap is thus filled with the appearance of the present work.

Professor Piggott is no stranger to the literature on the subject: he has been writing about prehistoric wheels since 1949 and in his wide range and detailed knowledge of the subject has no rival. As often happens, a number of recent books have dealt with aspects of the same problem, and here mention must be made of M.A. Littauer and J.H. Crouwel's *Wheeled Vehicles and Ridden Animals in the Ancient Near East* (Brill, 1979), Crouwel's *Chariots and other means of Land Transport in Bronze Age Greece* (Allard Pierson, 1981), and volumes in the series *Prahistorische Bronzefunde* dealing with the Italian vehicles (E. Woytowitsch) and early horse harnesses (H.-G. Hüttel).

But Piggott is concerned to write a connected narrative tracing the development of wheeled transport from earliest times

down to Classical Antiquity, and this he does in typically thorough and workmanlike fashion, apparently equally at home in the unfathomable awfulness of Russian excavation reports of 3rd millennium 'Pit Grave' vehicles and in the complexities of Iron Age wagon burials in central Europe. Thus we progress from early solid-wheeled wagons which would have been pushed to travel at walking pace, no doubt used for getting the hay in and the manure out, through the first spoked wheels as seen on Bronze Age chariots, to the marvellous and complicated carriage of Dejbjerg, with its roller bearings and pivoted front axle. The progression is not merely technological: the social effects of wheeled transport are in many ways the most important thing about it, as chieftains and other elite groups contrived to make their status manifest by riding, and eventually having themselves buried, in or with wagons and chariots.

In spite of the 'coffee-table' appearance this is a specialist's book which the layman may find hard going in parts. Careful study will however repay the effort, for this is a mine of information on many subjects. It is worthy of note that Piggott relies almost entirely on radio-carbon dates for the earlier periods, a position which will not find favour on the European continent. Nor will Near Eastern specialists care for: "The absolute chronology of Bronocice is more reliable than that of Uruk IV or the Early Dynastic periods of Mesopotamia".

One topic that I feel should have been more fully covered is Aegean Bronze Age chariots, but this is a relatively small failing in such a feast of learning and wisdom. The book is beautifully produced (in Hungary) and marks a worthy culmination to the author's long and painstaking work in such an important area of technology. □

A.F. Harding is a Lecturer in Archaeology at the University of Durham and a specialist on European Bronze Age archaeology.

New ideas — new science

J.Z. Fullmer

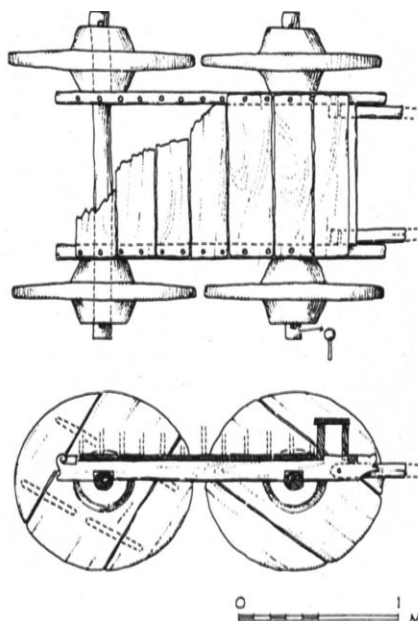
The Revolution in Science 1500-1750.

By A. Rupert Hall.
Longmans: 1983. Pp.373. £8.95.

IN 1954 Longmans published A. Rupert Hall's *The Scientific Revolution, 1500-1800*, a work so well received that 1962 saw a second edition. The book under review is a heavily revised third edition. Only two chapters, that on the influence of technology on scientific development, and that on 'The Range of Life', remain close to their 1962 version. Hall was moved to drastic revision — as reflected in the changed title — by his desire to mirror "the increased maturity of studies of sixteenth and seventeenth century science". To accommodate this new scholarship, Hall has lopped off the last half of the eighteenth century, along with discussion of the changes operating in chemistry and in the electrical sciences.

What scholarly studies moved Hall to change his emphasis? 'Notes' appended to each chapter provide clues toward an answer. For Galilean studies the work of Stillman Drake; for Newtonian studies, the work of D.T. Whiteside, of A.R. and M.B. Hall, of B.J.T. Dobbs, and of R.S. Westfall; for biological investigations, the work of Walter Pagel. The list could be extended from Hall's 'Notes' to the sections on the organization of science, on Copernicus, on Kepler, on Tycho Brahe, and the like. Still, the list, while understandable — for the work of these distinguished scholars certainly must be taken into account in any text on the scientific revolution — raises a serious question. Does the reader assume that changes in chemistry and in electrical studies are not part of 'the scientific revolution', or does the reader assume that scholarly definition of the term has been revised to exclude changes that appeared in these two disciplines? Hall does not address this question directly. Therefore, the reader familiar with the earlier edition is left with an impoderable; there is no way of knowing what has prompted the exclusion of these fields.

While some sections have been expanded (and others truncated) some new scholarship has been accommodated, and there is now a mountain of papers and books bearing on his topics and periods which Hall has neglected. What has remained constant is Hall's attitude and his range of view. He avers that he wrote about ideas rather than about society, and about particular individuals rather than about anonymous masses. Early in the text (p.2) he announces that he "unashamedly" follows "a positivist or even whiggish line". He asserts his belief in "progress"



Above: Remains of a wagon from the second millennium BC found at Trialeili, Georgian SSR.

Left: Scale drawing of a wagon from the same site.

and is convinced that "in the systems of scientific knowledge progress can be measured in terms quite independent of value judgments" (p.4). Newton's world-view, he argues, represents progress over Aristotle's because "Newtonian theory is bigger, more exact, more precisely testable, and above all more mathematical than its predecessor." Hall views the history of science as a branch of the history of ideas. Surprisingly, he equates the force of Newtonian ideas with the force of Lenin's ideas (p.360).

Hall has ignored, or rather, neglected to include in his treatise any accounts of the social relations of science, and little of the political relations. Martin Luther, for example, warrants mention only three times, and these mentions are incidental; it is not until page 348 that the reader is reminded that war separated the French and the English from 1698 to 1714. Hall's second chapter, 'The Problem of Cause' may be regarded as an attempt to deal with the issues raised by the sociologists of science, but most of the scholarship is dealt with indirectly, by refutation. The reader is told why Thomas Kuhn's *The Structure of Scientific Revolutions* (University of Chicago) should not be applied to the scientific revolution of 1500–1750 but Kuhn's work is not directly cited in this connection (citations to it do appear elsewhere). Merton's thesis about the role of Protestants in the founding of the Royal Society is rejected, although there is an acknowledgement that Christopher Hill and Charles Webster have found it seminal. Hall deals handsomely with Dame Frances Yates's magisterial *Giordano Bruno* (Routledge), but rejects its operating thesis. These are, of course, exactly the positions to be expected from someone who sees the history of science as a sub-set of the history of ideas.

In sum, Hall's book tacitly redefines the term 'scientific revolution' by concentrating almost exclusively on changes in astronomy, mechanics and biology, removing from consideration changes in chemistry and electricity; the scientific revolution is seen as a triumph of mathematicization, as the progress of rationality over religiosity, above all, as an 'internal' process depending almost exclusively on 'progress' within a discipline. There is no doubt in this reviewer's mind that if what is wanted is such an 'internal' history, Hall's book is one of the best of the genre. The prose is lucid, the structure of each chapter (and of each paragraph) is rhetorically exemplary and graceful. Graduate students, those with some knowledge of the literature, should find the book rewarding reading; undergraduates will be confused by it. Non-historians, especially those with 'internalist' or whiggish proclivities will find here stong confirmation of their own convictions. □

J.Z. Fullmer is a Professor of History at the Ohio State University, Columbus, Ohio.

The ascent of man

Robert Fox

The Montgolfier Brothers and the Invention of Aviation 1783–1784.

By Charles Coulston Gillispie.
Princeton University Press: 1983.
Pp. 210. \$45, £30.20.

THE first public demonstration of a hot-air balloon took place on 4 June 1783 at Annonay in France. The setting, in a remote, hilly part of what is now the department of the Ardèche, was an improbable one, and the curious onlookers who gathered in the town's square could hardly have appreciated the momentousness of a flight that lasted a mere ten minutes and ended frighteningly with the sackcloth and paper of the balloon's fabric set alight by the fire of its brazier. But within weeks the brothers who mounted the demonstration, Joseph and Étienne Montgolfier, had leapt from provincial obscurity to the rank of national celebrities, and they have remained celebrities ever since.

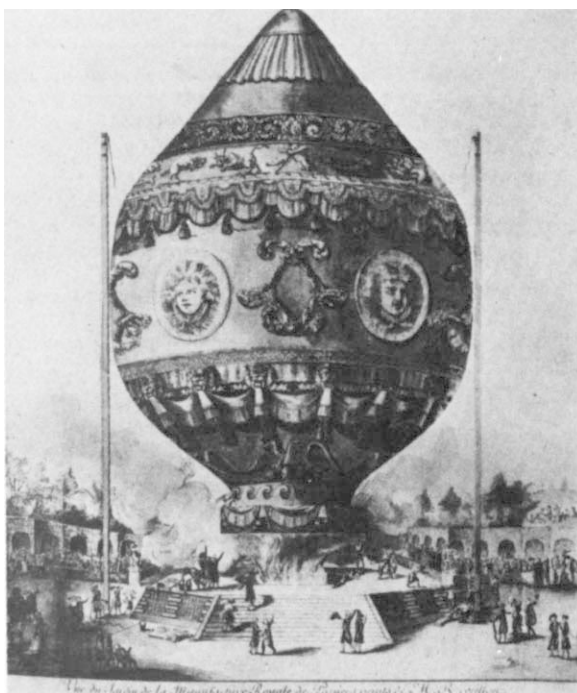
Charles Gillispie devotes the greater part of his gracefully written book to an account of the circumstances of the Montgolfiers' invention and of the early heroic years of aviation in France. It is a story bristling with dramatic events and colourful characters. From the start, spice was added to the episode by the initial, mistaken assumption of the Parisian *savants* that the Annonay balloon must have been raised by hydrogen. The result was a rivalry, which the Montgolfiers never sought, between the

hot-air *montgolfière* and the hydrogen balloon, or *charlière*, named after the physicist J.A.C. Charles, who was a notable pioneer of hydrogen ballooning.

The degree of popular interest was extraordinary, and grand public demonstrations became the order of the day. In September 1783, Étienne Montgolfier staged a particularly memorable flight for the king and queen at Versailles and on 21 November 1783 the inevitable next step was taken when a prominent scientific lecturer, Pilatre de Rozier, and an infantry officer, the marquis d'Arlandes, undertook the first manned flight in an untethered balloon. Ironically, it was Pilatre de Rozier who became the first victim of aviation in June 1785, when his hazardous attempt to use a hydrogen balloon and a hot-air balloon in tandem ended in an all too predictable disaster. But even that well-publicized tragedy did nothing to dampen enthusiasm. Ascents continued to attract large crowds, and the risks only added to the stature of professional aviators, several of whom earned a good living by their displays. The public's appetite for daring deeds was insatiable until well into the nineteenth century, and it is only surprising that accidents (such as the one that led to the death of the colourful Madame Blanchard as she injudiciously enlivened her ascent in a hydrogen balloon by letting off fireworks) were so few in number.

The conviction that ballooning quickly became an altogether too sensational affair is unavoidable. This was certainly the view of the Montgolfiers, whose hearts and principal means of livelihood remained in their paper-making business back in the Vivarais, and the high jinks and attractive iconography have been a snare for historians ever since.

But this book dispels the idea that early ballooning was nothing more than an unscientific activity divorced from all seriousness. The reinterpretation is an important one, resting on printed sources, letters, and notebooks that have been hitherto unread or inadequately exploited; and it appears unchallengeable. The theoretical content of the Montgolfiers' work may have been unconventional, but it was substantial and, in significant respects, original, as Gillispie demonstrates. He shows Étienne Montgolfier engaged in calculations of the lifting power of balloons ('the bigger the better' was the gist of his conclusions) and embarking on the design of paddles for a dirigible balloon. And, most interestingly, he points to



Testing an early balloon — October 1783.

Joseph's abiding curiosity about the relations between heat and mechanical work.

Clearly, it is hard to know just how far-seeing these wider reflexions were. Unlike his better-educated brother, Joseph was a doer, a self-taught inventor, rather than a *savant*. But it seems likely that from the 1780s until his death in 1810, he groped some way towards the notion that heat and work are interchangeable. This, at all events, was the retrospective view of the engineer Marc Seguin, a great-nephew of the Montgolfiers who, as a child, had known and learned from Joseph. Gillispie's careful re-examination of the evidence suggests that Joseph's ideas probably originated in his work on an ingenious heat-pump, a device that raised water by means of the heat produced in the rapid explosion of hydrogen or the burning of faggots. Thereafter, it survived through

a series of related inventions of which at least one, the hydraulic ram, was favourably viewed by the leading scientists of the capital. I find the historical significance of this work as intriguing as Gillispie obviously does, though it throws up tantalizing possibilities rather than a justification for a major reinterpretation of the prehistory of thermodynamics.

Despite the loose ends, it is hard to imagine that there is very much more to be said on the Montgolfiers and their world. However, quite apart from its scholarly merits, the book can be recommended as a good read. The work of any historian has its drab moments, but the wit and liveliness of Gillispie's text suggest that in this case such moments were few and far between. □

Robert Fox teaches the history of science at the University of Lancaster. At present, he holds a British Academy Readership in the Humanities.

Megalithic points

Clive Ruggles

Echoes of the Ancient Skies: The Astronomy of Lost Civilizations.

By E.C. Krupp.

Harper and Row: 1983. Pp.386. £16.95.

POPULAR books on archaeoastronomy continue to issue forth from the presses once a year or so, although the emphasis has changed distinctly since the 'Stonehenge Decoded' debate of the 1960's which started it all, and again since the more recent controversies over Alexander Thom's theories on megalithic 'science'. The current message is that astronomy is a sacred rather than a 'scientific' activity; and that in studying archaeoastronomy we are exploring the belief systems, ceremonial activities and cosmologies of diverse societies past (here we now include the great literate civilizations) and present (in this case, for 'archaeoastronomy' read 'ethnoastronomy').

The subject matter of Krupp's book (excluding what amounts to a rather incongruous swift potted history of cosmology in the final chapter in order, apparently, merely to make a concluding point about "why we do it") reflects a change in the definition of archaeoastronomy which is all to the good; archaeoastronomy is beginning to find its feet as a respectable branch of archaeology and ethnography. There is talk of how rituals serve to demarcate and regulate time periods (hardly a new point to anthropologists, admittedly) and hence tend to be related to astronomical observations; and of how the extent of astronomical practice can relate to a society's complexity. It is good to see such points stressed at intervals throughout *Echoes of the Ancient Skies*, and illustrated with a variety of examples.

Yet I feel, such success has been achieved

in the face of, rather than owing to, the thematic way in which the material has been organized. The Maya civilization (to take an extreme example) crops up in no less than six different chapters and might arguably have appeared in one or two more. On the other hand, in the 'Skies we Watch' chapter we are taken in the short space of 40 pages from ancient Egypt to Shang dynasty China, thence to Babylon, pre-conquest Illinois, prehistoric Scotland and Brittany, the Inca in Peru, the Maya in Mexico, and finally back to ancient China.

All this leaves the reader breathless if not a little confused and at times, as one apparently unrelated description follows another, perhaps even slightly bored; this despite Krupp's enthusiasm and readable style, and the originality and excitement of much of the subject matter. Attempts to draw threads together are too few and far between, and I wonder if, in the end, archaeoastronomy doesn't come over to the average reader primarily as the mere documentation of astronomical practice.

Krupp has to some extent played down controversy in order to present the evidence for ritual astronomy as a coherent whole. This is defensible for a popular book, but care is needed that a new popular bandwagon (albeit less fantastic than previous ones) does not start rolling.

Already ideas of ritual astronomy are running well ahead of the evidence in some areas, notably the megaliths (where serious debate continues, on the basis of extensive new site surveys and new methodology, about the precision and often the very existence of significant astronomical alignments). I feel that Krupp might have emphasized these areas: by doing so his book would have been truer to the spirit of much current serious archaeoastronomical research, and might have attracted a more discerning readership. □

Clive Ruggles is a Research Fellow in the Department of Mathematics at the University of Leicester.

Falling apples and leaning towers

John D. Barrow

Frame of the Universe:

A History of Physical Cosmology.

By Frank Durham and

Robert O. Purrington.

Columbia University Press: 1983.

Pp.275. \$32.50, £22.

'HISTORY repeats itself; historians repeat each other'; the natural reaction perhaps of someone seeing another medium-brow guide to the history of science. A sufficient number of such surveys already exist to create a form of '1066 and all that' sub-culture within the subject. Those events that never happened live most vividly in the mind and are much more memorable than those that really did: Newton's falling apple; Archimedes' bath, Galileo and the tower of Pisa — all are cornerstones of the sub-culture.

The authors of the 'Frame of the Universe' have chosen to follow the history of cosmological ideas from ancient to modern times at a level suitable for non-specialist college courses and general reading. They claim no great novelty in their treatment and erect their frame around the first ancient and Greek astronomers, the Medievals, Copernicus, Galileo and Newton, before moving on to Einstein, modern big bang cosmology and gravitational collapse. Yet, the clarity of presentation and the engaging style of the authors make this an enjoyable book for any scientist to read. Those wishing to pursue subjects in greater depth are provided with an excellent bibliography and detailed notes.

One of the problems with histories of this sort is that our own categories of thought so influence the presentation. We view the past solely in terms of the route necessary to reach the (right) answers of the present. The failures are ignored as inessential by-products of a never-faltering march towards the 'truth'. This 'Whig' approach to the history of science is the one that prevails in the minds of most working scientists with a passing interest in the history of their subject and, although the authors are aware of this snare, they do not make any real effort to avoid it. The other weak point in the overall treatment, which will be disappointing to many students reading the book, is that whereas the authors are very lucid in describing the course of events, they are weak on the explanatory side. They rarely ever ask the interesting question 'Why'. Why, for example, did the Jews take no interest in astrology? What role did their religious beliefs play, and so forth.

In the opening chapters there is a particularly clear discussion of various megalithic 'observatories', including Stone-

henge. Here, and throughout the authors make good use of clear diagrams and some simple geometry to explain details. In the early chapters they do, to my mind, make too much of the idea that most of the content of related archaic myths is induced by astronomical experiences since these were the only universal experiences of early peoples. If we look at the laws of ancient societies we also find an unexpected homogeneity, yet no one would claim that this was astronomically induced.

The ancients lead inevitably to the Greeks and they to the Medievals. Medieval Science was but Greek science written in Latin. Unfortunately, the non-European cultures are all but excluded from the description; nor do the authors ever dwell on the interesting question as to why science developed so effectively in Europe whilst the early Eastern and Oriental successes faded. Some treatment of non-European scientific development would have been valuable, if only because there have recently been a number of ridiculous publications about the relationship of modern physics to Eastern religions.

The Medievals lead inevitably to chapters on Copernicus and Galileo. We see the end of the medieval mindset: the authority of texts gives way to the authority of observations. But, the new dogma, like the old, also has its problems if applied inflexibly, for Copernicus had no concept of the inevitable experimental error associated with any observation or measurement. The authors show how this blindspot influenced Copernicus' deductions and point out that even Newton and his contemporaries failed to appreciate the relation of theoretical predictions to actual measurements. The Copernican development is a good illustration of the 'Whig' approach, as it always appears as the turning point in the history of astronomy. In *retrospect*, of course, it is; what is now equally clear is that *De Revolutionibus orbium coelestium* had negligible influence until the seventeenth century. Few copies of it were sold, and even fewer read in the early years after Copernicus' death. Other great events, like the Portuguese voyages of discovery, entirely overshadowed it.

The next stop is Newton and the *Principia*. After this there is a leap into modern times. General relativity, the big bang and gravitational collapse make up the last quarter of the book. Although the treatment of these subjects is accurate it is extremely lightweight. More mathematical detail was devoted to explaining Ptolemaic epicycles than special relativity. Also, the brief description of elementary particles and the role of symmetry principles was too brief to be of any value.

Another gap in the development was to ignore the way theories were evaluated and the relationship of cosmology to philosophy. In ancient times these two lines of thought regarding the 'frame of the

universe' were inseparable but after Descartes no philosopher made a significant contribution to science. The modern era sees a divide between the two disciplines, each of which cultivate different interests. A discussion of how this divide arose and what now distinguishes the two approaches would have been valuable. Also, the authors ignore discussing another modern development: how science thinks about itself. Most scientists take as their article of faith some version of Popper's naive falsification as a guiding principle. However, others following Duhem have stressed that hypotheses can only be tested in bundles, never in isolation. We can save any cherished

hypothesis from falsification indefinitely by invariably picking as false another member of the 'bundle' involved in the experiment. In fact, the way in which the false assumption is picked from the 'bundle' is largely what separates 'crank' science from real science.

This was a book I enjoyed reading and one which undergraduates in historical subjects should find useful as an introductory guide. British science students suffering from educational overspecialization might also enjoy seeing when and how authority changed. □

John D. Barrow is a Lecturer in Astronomy at the University of Sussex.

Celestial talk

A. Rupert Hall

Telescopes, Tides, and Tactics:

A Galilean Dialogue about the *Starry Messenger* and Systems of the World.

By Stillman Drake.

University of Chicago Press: 1983.

Pp.236. \$22.50, £18.

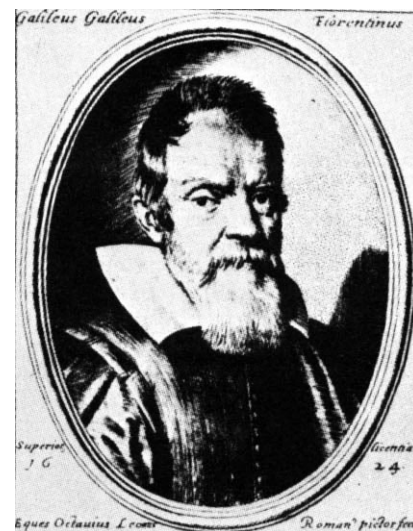
THE translator of all the genuine Galilean dialogues, Stillman Drake, after presenting his version of Galileo's early work in physics in the form of an imaginary dialogue between three of Galileo's friends under the title *Cause, Experiment and Science*, now employs the same vehicle to describe the first few years of Galileo's activity as an astronomer.

The scene is Venice, the year 1613. Galileo is in Florence. The learned talk is between Sagredo and Salviati — two friends whom Galileo himself used as expositors — and Paolo Sarpi, a religious and a philosopher, another real friend of Galileo's introduced here because Galileo first outlined his theory of the tides to Sarpi (in 1595) and first stated to him the law of fall (in 1604). The talk, without 'Prithees' and 'Forsooths', is in the modern language of a translator, in a style (to my ear) rather more formal and stilted than that actually employed by Galileo himself. Perhaps as these are imitations of imaginary conversations, that is right. Whether the dialogue form is preferable to straightforward historical exposition is a matter of taste; the facts and arguments are clear enough here, and Drake provides references for those sentences that have been lifted more or less literally from Galileo's own writings.

A substantial part of the book is the 'reading' of Galileo's *Starry Messenger* (*Sidereus Nuncius*, 1610), of which Drake published a partial English version many years ago. The printing of a complete translation now, with all the details of Galileo's observations of Jupiter's satellites, enables Drake to give an analysis — part historical, part reconstruction — of Galileo's longest

continued, most exact and most frustrating investigation in observational astronomy, which never yielded that practical solution to the problem of longitude at sea, ardently hoped for by the discoverer. This analysis, in turn, adds further substance to Drake's view of Galileo as, above all, an empirical investigator of nature.

There are many other issues raised in the dialogue, not all of them firmly resolved (it is a real advantage of the conversational vehicle that it allows a degree of uncertainty common in life but rare in historians). It starts from the question: why



Galileo Galilei

has Galileo not published the *System of the World* announced in print in 1610 and in private even earlier? This, together with the related theme of the mechanical explanation of the tides, permits Drake to put forward new ideas about Galileo's development as an exponent of the new astronomy — in which context his relationship with Kepler, another theme, is also relevant. In particular Drake introduces the plausible hypothesis that Galileo — whom history knows only as a student of motion for his first thirty years and more — was led to take Copernicanism seriously by hitting upon his 'inertial' explanation of the tides (1595). This explanation was,

Drake argues, central to Galileo's astronomical system, though it was pushed somewhat into the background by papal pressure before the publication of the *Dialogo* of 1632.

Drake does not here defend the scientific adequacy of Galileo's theory of tides, nor the adequacy of his concept of inertia, nor one or two other positions which have seemed idiosyncratic to other scholars. His 'reconstructions' are balanced, clever, and well-documented. As in other books by him, 'the philosophers' come in for a good many hard knocks: 'There is a vast difference between confidence grounded in repeated observations and careful thought, and the confidence of philosophers in metaphysical principles that transcend all experience of mere astronomers', remarks Salviati. Drake (and other scholars) have successfully vindicated their claim for the historical veracity of Galileo's experimental approach to mechanics; here he extends the lesson to the successful

matching of mathematical formulations with carefully compiled observations on Galileo's part. There is a great deal here which is new and really important. Drake stands with those who argue that the historical process of the separation of natural science from philosophy, metaphysics and, indeed, magic was just as important as the fact of the continuity of elements derived from these sources in natural science. For it was by separation that the special character of science was established.

Drake has tried to give a new and more lively form to the often pedantic literature of the history of science, and so to arouse a new curiosity. Whether or not he succeeds in that aim, his imaginary conversations, like those of Galileo himself, contain new truths vividly argued. □

A. Rupert Hall is Emeritus Professor of the History of Science, Imperial College, University of London.

Stepping through the Universe

A.T. Winfree

Powers of Ten: About the Relative Size of Things in the Universe.

By Philip Morrison and Phylis Morrison and the Office of Charles and Ray Eames.

*W.H. Freeman: 1982. Pp.150.
\$29.95, £11.95.*

TWELVE months ago, the editors of *Scientific American* launched the Scientific American Library as a book club, with books appearing at two month intervals, each being reissued as a trade book 12 months later at a higher price. Intended for a popular audience, the series is pitched at a level somewhere between the monthly magazine articles (which I personally consider demanding reading) and *Time/Life* picture books. Each is written by a scientist, many by Nobel Laureates, not by the editorial staff. Each presents a subject that cuts across diverse specialized disciplines. The first six issues have already been produced.

Powers of Ten is the premier issue and the first available to the general public. It is by far the least technical of the books produced up to now, and demonstrates the publishers' commitment to readability and quality graphics, in 22 cm square format with colour pictures on almost every page. The theme is familiar and not subtle: phenomena in the natural world vary over something like 42 orders of magnitude in linear dimensions (ignoring volume, time scales, etc.), and differ in character according to scale. Many books have presented this theme to diverse audiences. So in this case everything depends on

targeting the right audience and on skill in presentation. The Morrisons' art work is superb, as one might have expected from their track record, and from the immediate antecedent of the book: the wonderfully successful ten minute motion picture of 1968 and 1977 (improved) by the same name.

The core of *Powers of Ten* is a sequence of 42 colour plates, the first suggesting most of the Universe. In the next plate (10²⁴ metres) some galactic clusters become big enough to see, as a giant step forward brings us ten times closer to the nucleus of a certain carbon atom in the very centre of that cosmic field of view . . . and so on through 42 steps. By the 25th step we have converged on a couple lying on a picnic blanket and we continue, ending after 17 more steps in a subnuclear blur that might represent quarks.

Facing the colour plate at each step is a page of text and smaller pictures illustrating some of the phenomena characteristic of that size scale. For example facing the 10 metre to 10⁻⁶ metre pictures, most of the illustrations are biological. On every scale, the chosen examples are rich with surprising tidbits of insight, presented from a unique perspective.

This core is only 60% of the book. The appendices are no less fascinating; in fact the one entitled "Rules for a Journey of Mind and Eye" might be the best place to start: it tells how the pictures were planned. My own taste is then to run backwards through the plates, starting with quark-blur. Other appendices gather some pertinent history of science; the 20 pages of "Sources and Notes" make particularly good reading. □

A. T. Winfree is Professor of Biology at Purdue University, Indiana.

ADVERTISEMENT

The Diversity of Green Plants

Third Edition

Peter Bell and Christopher Woodcock

Contemporary Biology Series

A well-illustrated evolutionary account of the photosynthetic plants, from the algae to the flowering plants, paying particular attention to those plants prominent in current research.

£9.95 paper 368 pages

The Biology of Lichens

Third Edition

Mason E. Hale

Contemporary Biology Series

A review of all aspects of lichenology.

This new edition has been rewritten, stressing ultrastructure studies and taking account of the tremendous progress in the study of lichen resynthesis and the physiology of symbiosis.

£8.95 paper 200 pages

The Physiology of Flowering Plants

Third Edition

H.E. Street and Helgi Opik

Contemporary Biology Series

This book treats plant physiology from the viewpoint of the whole plant or organ, with particular emphasis on growth and development. New additions to this edition include a treatment of the general physiology of photosynthesis, and an expansion of the topic of stress resistance.

£8.50 approx paper 296 pages January

Immunity to Parasites

How animals control parasite infections

Derek Wakelin

The book concentrates on selected host-parasite systems, all of clinical or economic significance, which illustrate particular aspects of parasite adaptation and of host response, and in which immunologically oriented research has made significant progress.

£9.50 approx paper 192 pages March

The Biology of the Coccidia

Edited by Peter L. Long

This is the thoroughly revised and updated version of *The Coccidia* edited by the late Dr Hammond and Dr Long. It includes new chapters on chemotherapy and new details of life cycles of several species of Sarcocystis and Toxoplasma.

... this has turned out to be a significant volume which in the present climate of scientific work may have been produced just in time.

British Veterinary Journal

£45 boards 512 pages



Edward Arnold

41 Bedford Square
London WC1B 3DQ

Circle No.35 on Reader Service Card.

The Universe as a Whole

P.B. Medawar

The Return to Cosmology: Postmodern Science and the Theology of Nature.

By Stephen Toulmin.

University of California Press: 1983.

Pp.283. £17, \$25.95.

WHEN I was a student the scientists most prominent in the public eye were theoretical physicists with cosmological leanings, men whose views on God and the order of the Universe were eagerly sought and widely attended to.

Whether or not they were the patricians of science they were reputed to be, they had a lot going for them; not the least important characteristic being natural good manners that would have made it unthinkable for them to make pronouncements upon, for example, a supposedly extraterrestrial origin of life. Arthur Eddington, moreover, wrote with a winning charm that has not been surpassed in popular scientific writing and Sir James Jeans had the courage to think big — to think about God and about the relevance to the garden as a whole of the patch of ground he himself cultivated. They had philosophic opinions too ("the stuff of the world is mind-stuff" declared one of them).

ADVERTISEMENT

"OUTSTANDING"

THE ATOMIC COMPLEX

by BERTRAND GOLDSCHMIDT

An accurate, complete and fascinating worldwide political history and personal memoir of nuclear energy — from the development of the bomb in World War II to today's nuclear energy complex and proliferation problems.

"READS EASILY"

Goldschmidt, a leading French scientist turned international statesman, reviews half a century of political moves, countermoves, international intrigue and manipulation. *The Atomic Complex* is a pragmatic look at the nuclear world today, carefully examining the terror nuclear weapons represent, but at the same time stressing the benefits of nuclear energy.

Order this 500 page
best seller today.

\$31 Hardbound \$24 Softbound

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Their principal weakness was to be hopelessly inexpert as philosophers: Eddington's epistemology was a pain and both he and Jeans were thought to have deserved the sharp repudiation of their opinions by Professor L. Susan Stebbing in her *Philosophy and the Physicists*.

Stephen Toulmin raises here anew that which might be called "Kant's problem":

Is "the Universe as a Whole," so considered, a legitimate topic of rational investigation and speculation, or even a legitimate subject of human language at all? . . . In what respects, and on what conditions, can anything be said about the natural world in its entirety which is not dependent on our ability to subdivide natural phenomena into separate aspects, along disciplinary lines, and discover truths about those aspects one at a time?"

I know no one better qualified to address these problems than Stephen Toulmin, a pupil of Ludwig Wittgenstein's. The way he goes about it is this: a long introductory chapter discusses Kant's problem as a whole ("What kind of things. . . can we learn from the results of the natural sciences which are of cosmological significance today?"). There are some amusing asides in the course of this chapter, in which he recalls his review of Julian Huxley's notion of evolutionary ethics under the title "World-stuff and Nonsense". He comments on Pierre Duhem's dilemma that it arose out of his unwillingness as a physicist to express any opinion conflicting with his personal commitments as a Roman Catholic, and indeed Duhem reserved to religion any right to speak about Reality. Of Teilhard de Chardin he charitably remarks that "his whole charming scheme of 'noosphere,' 'omega point,' and the rest, still strikes me as being a vehicle of wish fulfillment as much as of serious thought".

Before he gets down to cases — to a consideration of individual cosmologists — Toulmin gives us an especially interesting and well-thought-out chapter on scientific theories and scientific myths. No one need fear that a man as sensible as Toulmin is going to propound the view heard sometimes from anthropologists that myths and scientific theories are alternative explanations for the world of the same stature, though one of them, myth, is emancipated from the bondage of a commitment to empirical truth. In my opinion, to compare science with myth is to compare the wrong genres: it is not myth and science but myth and imaginative literature that go together, for what are most myths but science fiction of forgotten origins? This is not quite the view of Stephen Toulmin, who makes a case for saying that mythical constructions may enter into professedly scientific theories to gratify the personal predilections of scientists to fulfil various non-scientific motives. Laymen will often be unaware that scientists are playing these tricks on them, because the habit of consulting scientists upon scientific problems to which they often have the answers tends

to invest scientists with a sacerdotal aura — with the effect that laymen look to scientists for authoritative opinions on matters on which they are entitled no more than private opinions.

The later part of Toulmin's book is devoted to a consideration of individual cosmologists — in particular of Arthur Koestler, Pierre Teilhard de Chardin, Jacques Monod, François Jacob, Carl Sagan and Gregory Bateson. The essays here are critical, to be sure, but they are unfailingly generous and well-mannered, so anyone on the lookout for good old-fashioned knocking copy will be disappointed. I should like to quote the opening paragraphs of Toulmin's piece on Koestler because they illustrate better than any other passages Toulmin's style of thinking and writing:

This is a bad time for polymaths. The old jibe about a Jack-of-all-trades being master of none has bitten deep into our minds, so that few people will admit to an intellectual grasp of anything more than a narrow range of experience. In a fragmented culture, everybody is *expected* to be a specialist: so men cling to the professional standards of their guilds as the lifebelts which will keep them afloat on a sea of general ideas which they have lost the capacity either to swim in or to plumb. . . . Natural scientists, again, look with suspicion at those colleagues who stray too far outside their specialisms, and for the most part they shut their eyes to the very existence of philosophy. In return, the philosophers have made their craft a 'profession' of its own: a professional philosopher no longer needs to understand even the broadest ideas of contemporary natural science — to say nothing of its factual discoveries.

Toulmin and I in our writings have covered much the same ground and not unexpectedly I don't always agree with him. I don't agree with Toulmin's generous appraisal of Arthur Koestler because I was more oppressed than Toulmin was by the element of "quackery" and show-offery for which Toulmin soundly castigates Koestler on p.139. Koestler had also a deep-seated lack of humility — made apparent by his campaign of propaganda (fortunately fruitless) against legislation that requires pets that might import rabies into the country to go through a period of quarantine before admission.

The best of Stephen Toulmin's very good writing is on Teilhard de Chardin, on Jacques Monod and — the deeper thinker in Toulmin's view — on François Jacob. Although he is sometimes sharply critical, particularly of Monod, he comes sometimes to the defence of the subject of his criticism — a defence justified by his perceptive understanding of the French cultural tradition, especially the element of exaggeration and the theatrical in French cultural dialogue. *Les anglophones*, he tells us, may easily be mistaken to "read their debating points as assertions, their exaggerations as dogmas, and their oratory as rigorous proof".

Writing of the eighteenth century, Toulmin comments:

If we are puzzled by the shelves of collected sermons in our ancestors' libraries, that is because we forget how far scientific and aesthetic questions have replaced moral and theological ones as the staple of dinner-table-talk; and how far the popular scientists has won over the audience of the popular preacher.

In Johnson's day conversation turned easily and naturally to ethics, philosophy or theology, and acquaintance with the writings of Andrew Baxter and Bishop Wolverton was taken for granted. Toulmin deplores as I do the change of fashion to-day that has withdrawn these topics from the dialogue and serious reading of so many educated people; to be sure it would

cast a pall over almost any modern dinner party if a hostess were to start the conversation at dinner by turning to her principal guest and saying "Pray, Sir Alfred, what is your conception of the notion of convenient Grace?"

This change of manners will unhappily deny Toulmin's book the readership that his book richly deserves because it is authoritative, especially well written and immensely readable. □

Sir Peter Medawar is a member of the Scientific Staff of the Medical Research Council's Clinical Research Centre, Harrow, Middlesex.

Maturing astronomy

G.T. Bath

Cosmology and Astrophysics: Essays in Honor of Thomas Gold.

Edited by Yervant Terzian and Elizabeth M. Bilson.

Cornell University Press: 1983. Pp. 163. £19, \$29.95.

"THAT Trinity of angry young astronomers of the fifties", was the way I always thought of them — Bondi, Gold and Hoyle. A trio of *enfants terribles* raging around the ivory towers of Cambridge, upsetting the established order of things, and putting God in his place by advocating a better way of ordering the Universe than He had been smart enough to dream of.

Of the three, Hoyle was the most articulate and arresting to the general public. Bondi was the distant mathematical genius with the strong foreign accent, the professor out of Tintin comics. Gold was the force in the background, thoughtful and taking great care to be patiently logical in his explanations of the way things ought to be on the basis — the only acceptable basis — of straightforward physical reasoning. Now, 35 years after the birth of steady state cosmology, the angry young men have become grand old ones. And we give them our honours.

This book contains essays in honour of Tommy Gold, celebrating his sixtieth birthday and covering the three main fields his sparkling mind has contributed to — cosmology, pulsar models and Solar System physics. They are written by past collaborators, including Hoyle and Bondi, and scientists of a younger generation who have been inspired by his person and by his ideas. That inspiration comes through clearly in articles by former students like Peter Goldreich — inspiration derived from outspoken courage and innovative cussedness. A succinct image is provided by Goldreich: "In 1963, I was a medal winner at the NCAA judo championships. Shortly thereafter, Tommy expressed doubt about the efficiency of judo. We

were standing in the kitchen at the time and I strangled him on the spot".

The first half of the book contains articles by Hoyle and Bondi on steady state cosmology, with a short historical memoir from Gold. The origins of steady state cosmology in 1948 make fascinating reading. It becomes clear how the development of different facets of the theory depended, in practice, on the interplay of personalities around the key concept of the Perfect Cosmological Principle, throwing out conservation of matter and energy with gay abandon, or modifying Einstein's equations to create a reservoir of available energy. The differences between the original ideas of Bondi and Gold on the one hand and of Hoyle on the other are made clear.

Hermann Bondi's contribution consists of a lucid exposition of the need for a general theory of relativity using the sort of gedanken experiment with which a scientist in command of his subject can explain the most abstruse and intuitively paradoxical aspects of physical law. Using chains, mirrors, buckets and cylinders, Bondi builds before our eyes the gravitational red-shift and Einstein's theory of gravitation, with illuminating asides as the façade rises. Hoyle's contribution attempts to resuscitate the cadaverous remains of steady state

cosmology. The greatest contribution of steady state cosmology, and, in part, of Hoyle's greater work on stellar structure and nucleosynthesis, was to inspire whole generations of younger astrophysicists. What greater contribution is there? It is their view which cosmology now reflects, with its close links with elementary particle physics, grand unified field theory and quantum gravity. It would sadly seem that Sir Fred has not been noticing.

Dennis Sciama contributes a piece on black hole explosions which amply illustrates the state of mental excitement and joy that comes from the appreciation of new unities within the body of physics — connections between apparently distant limbs which, when recognized, can act to open up fresh channels of understanding. Of all the contributions to the book, I most enjoyed the short, sharp chapter by Joseph Taylor describing the present state of the field which he has done so much to create, the binary pulsars. The description is a classic of succinct, objective reporting, ending with a flourish which would have caused many a raised eyebrow in the not-too-distant past: "We take this agreement" (between the predicted and observed period decrease of the binary pulsar PSR 1913+16) "as compelling evidence for the existence of gravitational waves".

The book ends with a discussion of theories of the Solar System, some of which are still controversial while others have become accepted in the established body of planetary science.

The mixture of heresy and established physics, iconoclasm and accepted orthodoxy reflects Gold's own contribution to science. Tommy Gold, for all his gadfly ways, knows one element which is essential in good science. Old men's theories must satisfy the critical appraisal of the coming generation. Otherwise they will suffer the wrath of angry young men. □

G.T. Bath is a Research Fellow of Wolfson College, University of Oxford.



"That Trinity of angry young astronomers of the fifties" — Thomas Gold, Herman Bondi and Fred Hoyle at a conference in Berkeley in 1961.

Hammering out the meaning of strata

Derek Ager

Great Geological Controversies.

By A. Hallam.

Oxford University Press: 1983. Pp. 182.

Hbk £15, \$35; pbk £7.95, \$14.95.

POLITICIANS often complain that the media only see politics as a matter of fights between personalities, but nothing attracts public attention more than a good hard-hitting row. The same is true of science, especially an inexact one such as geology. Certainly one of the best ways to establish one's name in the subject is to be engaged in a hearty controversy. For all the great work done by Sedgwick and Murchison on the Cambrian and Silurian Systems respectively, nothing is so much remembered as their battle over the boundary between them. Do we not perhaps spend too much time on the disputes?

Tony Hallam's latest book is a learned and thorough discussion of several of the most famous geological controversies, at least in Britain. So he takes us through five "case histories": neptunism — vulcanism, catastrophism — uniformitarians, the Ice Age, the age of the earth and continental drift. He finishes with "general considerations" and an excursion into the philosophy of science. Hallam is unfortunate in that the publication of his book coincides with that on *Geology in the Nineteenth Century* (Cornell University Press) by M. T. Greene. Both books cover much the same ground, but in surprisingly different ways. Thus Green is much more preoccupied with the arguments on the European continent on the mode of formation of the Alps, which is hardly mentioned here.

Hallam divides the geological world into the arm-wavers and the nit-pickers: the great generalizers who stand on mountain tops (at least in spirit) and wave their arms about and those who keep their noses and their hammers close to the rocks. The former naturally get the glory and appear in books such as this, but it is the latter who do all the work and whose observations must ultimately be explained. That is why I prefer the views of Cuvier, who recorded what really happened bed by bed in the Tertiary rocks of the Paris Basin to those of Lyell who did not really look at strata in detail at all. Yet it was the latter who, in his day, triumphed and established for himself the status of the great thinker and synthesizer of geology, whilst the former was derided for his extremism.

It is noteworthy in this connection that Hallam proudly claims Cambridge as the leading centre of things geological in the years leading up to plate tectonics. This is

true if one only thinks of the geophysical side of Cambridge (before the three departments there merged). But the Geology Department at Cambridge was at that time still plodding away with the meticulous description of Lower Palaeozoic stratigraphy, as it had been for 150 years before.

Similarly, it is worth noting, as Hallam does, that "the prime goal of the founders of the Geological Society [of London] in 1807 was to eschew argument and speculation [as was then going on so actively in Edinburgh] in favour of sober fact-finding". This is not to decry the desirability for grand theorizing but it reminds us that we also need the careful, and many would say the only, real geology.

One theme that emerges clearly from this book is the nationalism of geological controversy. I well remember teaching continental drift to graduate classes in America in the late fifties. They were sceptical, as Hallam indicates was the general atmosphere in the US at that time, but they were not beyond persuasion even before plate tectonics. I think this illustrates the point that in many of these controversies, national, social and political factors played a big part in the scientific attitudes adopted in different countries. It will surprise many prejudiced people, no doubt to read Hallam's assertion that

supposedly autocratic and dogmatic German academia was more amenable to new ideas than most other countries. Certainly Protestant Britain went through far more agonies over the Noachian deluge (and later over evolution) than did her Catholic contemporaries. Also our monarchic, aristocratic background increased our fears of revolutionary ideas, both political and scientific, from the continent.

Apart from his fondness for words such as heuristic, Hallam's book is very easy to read. It may follow well-travelled paths, but it does so with a remarkable broad knowledge of the literature. It is full of lengthy but very relevant quotations. One may quibble at times (I wonder, for example, at what Norwegians would think of his statement that Switzerland is the only European country with substantial mountain glaciers) but essentially it is reasonable and well-argued. It presents a careful examination of contrasting views rather than the over-simplified versions we usually find in the literature. Certainly it fully respects the many truths contained within the apparently diametrically opposed opinions. □

Derek Ager is Professor and Head of the Department of Geology at the University College of Swansea.

A life of classical physics

John J. Roche

Ludwig Boltzmann. Man. Physicist. Philosopher.

By Engelbert Broda.

(translated by L. Gay and the author)

Ox Bow Press: 1983. Pp. 169. \$22.50.

THIS brief biography of Boltzmann was a delight to review because of its captivating style, and because of the mirth it frequently provoked. Professor Broda is a professional scientist well known for his writings on the history and philosophy of science. His text, which is a revised translation of a German text first published in 1955, is written in a relaxed and good-humoured manner and conducts us on a brief and somewhat personal survey of the life and times of Boltzmann.

Boltzmann is his hero, "one of the greatest thinkers of all nations and all times". Broda does, nevertheless, engage in a gentle and oblique criticism, which is quite penetrating and revealing. He has gathered together a fascinating collection of vignettes of Boltzmann, culled from friends, family and colleagues. A valuable aspect of this biography is that Boltzmann is largely allowed to speak for himself through a wide selection of well-translated passages. A portrait emerges of a very

attractive and sensitive man, a keen pianist, rather unworldly, someone who enjoyed the theatre, liked to travel, and had a considerable penchant for doggerel. Boltzmann was an excellent and humorous teacher, and most of his published books emerged from his teaching. In the 1890s Boltzmann was drawn into a heated and protracted controversy over the positivist opposition to the atomic theory which cast a shadow over his life's work.

Briefly, Boltzmann was born in 1844 in Vienna into a comfortable middle class Austrian family. He studied physics at the University of Vienna, where he received his doctorate in 1866. He worked at different times with Stefan, Loschmidt, Bunsen, Kirchhoff and Helmholtz. He spent six years at various German universities, but the rest of his academic appointments were in Austria. He does not appear to have enjoyed Prussian severity. Boltzmann married Henrietta von Aigentler in 1876, who bore him four children. He committed suicide during an attack of depression while on a summer vacation near Trieste in 1906.

Boltzmann was essentially a late nineteenth century classical physicist. Although he was witness to the birth of a new physics, he virtually ignored it. Boltzmann's work in physics was dominated by a commitment to mechanism and to atomism. In common with a great many physicists of the nineteenth century he combined the skills of an experimentalist with those of a mathematical physicist.

Boltzmann's best known contributions to physics lie in the realm of the kinetic theory of gases and in statistical thermodynamics, which occupied him for forty years. He extended the work of Maxwell and others, generalizing the analysis of the distribution of energy among atoms and molecules. Using similar methods he also generalized the analysis of transport



Ludwig Boltzmann, as drawn by his student Karl Przibran, later Professor of Experimental Physics at Vienna University.

phenomena in non-equilibrium systems, arriving at the famous Boltzmann transport equation. His greatest achievement lay in the discovery of a functional law which related thermodynamic entropy to the statistical distribution of an ensemble of molecules. This law, which was written by Planck in the form $S = k \log W$, is inscribed on Boltzmann's tombstone. Boltzmann also extended the mathematical analysis of entropy to non-equilibrium states which are not covered by the thermodynamic definition. His experimental work included the study of dielectrics and of diamagnetism.

In the late nineteenth century many physicists became impatient with the penetration of metaphysical concepts into physics, and with the din of competing theories in electromagnetism and in thermodynamics. There were two distinguishable responses to this predicament, positivism, led by Ernst Mach, and conventionalism, led by Boltzmann. Positivism dismissed hypotheses of all sorts, metaphysical and physical, as harmful, or as mere mnemonics at best and wished to reduce the laws of physics to a summary of observation-statement.

Largely because of the fierce positivist opposition to the atomic theory, Boltzmann was forced to articulate an alternative epistemology of physics, which combined the ancient mathematical device

of 'saving the appearances' with a dash of Kantian idealism. According to Boltzmann, hypotheses are indeed useful but not as candidates for physical 'reality', as we have no access to such reality, but rather as convenient pictures which stimulate and guide research and which can be made increasingly self-consistent. In this manner, at one stroke, Boltzmann undermined the pretensions of metaphysics and of theorists, an defended his right to an atomic 'picture'.

Positivism and conventionalism have competed for the minds if not the hearts of physicists ever since. Conventionalism, for example, is behind the reduction of force, momentum and the electromagnetic field intensities to mathematical constructs, and behind the reduction of Newton's laws to definitions.

It is also responsible for the contemporary preference for using the term 'model' rather than 'theory'. There were many

other sides to Boltzmann's philosophy, of course. He was an ardent Darwinian, for example.

Broda laments Boltzmann's lapses of literary style. Indeed, Boltzmann's historical errors, confusing Roger Bacon with Francis Bacon, for example; his description of the Greeks as 'naive' and his definition of metaphysics as 'mental migraine' heralded, perhaps, a decay of style in physics. Faraday, Maxwell or Helmholtz would not have committed such solecisms.

Broda's analysis of Boltzmann's science and philosophy does not represent exacting scholarship, and the translation has peculiarities such as "atomistics" and the "obligate increase in entropy". The non-specialist, however, will find this a good and informative read. □

John J. Roche teaches History of Science at Linacre College, Oxford.

Intelligent steps

Simon Lavington

Towards Fifth-Generation Computers.

By G.L. Simons.

National Computing Centre: 1983.

Pp.225. £10.50, \$18.10.

The Fifth Generation:

Artificial Intelligence and Japan's Computer Challenge to the World.

By Edward A. Feigenbaum and

Pamela McCorduck.

Addison Wesley/Michael Joseph: 1983.

Pp.288. \$15.55, £9.95.

FEW TOPICS in computing excite so much emotion as Artificial Intelligence. Since anyone possessing a brain — artificial or not — is entitled to join the debate, the task of disentangling fantasy from reality is not easy. The dream of artificial intelligence has been with us from the early days of computers, with several respected pioneers expressing the cautious hope that machines would one day exhibit intelligent behaviour. The reality has yet to appear.

Artificial intelligence research concentrated for many years on identifying and reproducing basic mechanisms of learning and problem-solving: laudable aims, but difficult to achieve with the primitive computing tools available and the fragmented nature of the research effort. General theories were slow to evolve, and the few demonstrations of problem-solving to emerge from computer science laboratories hardly seemed relevant to real-life applications.

A shift of emphasis took place in the 1970s, when some of the general techniques of reasoning, notably inference mechanisms, were applied to collections of specific knowledge. The aim was to make the computer appear as an expert assistant. Predigested knowledge, in the form of

rules and facts about, say, respiratory diseases, was embedded in a friendly question-answering framework and the so-called Expert System was born. Here at last was a money-making product which hard-nosed industrialists could appreciate. More importantly, there arose the conviction amongst a small but growing band of computer users that information systems could be made 'smart': that dumb database systems could one day be replaced by intelligent knowledge-based systems.

This vision of intelligent knowledge-based systems has not only become the acceptable face of artificial intelligence; it is seen by many as the essential ingredient of future computers. Nowhere is this more firmly believed than in Japan. The Japanese, casting around in 1979 for a policy for long-term economic survival, launched an investigation into the most likely trends for the next generation of computers. The resulting report described their image of the 'fifth generation' computer, and stressed the importance of three functions: problem-solving and inference-making, knowledge-base management, and intelligent interfaces. Soon the 'fifth generation' had become a peg on which anyone felt free to hang opinions and prejudices, hopes and fears, about every aspect of future computer hardware and software.

Although many strands of research were naturally assumed to contribute to the fifth generation, Japan's uncompromising inclusion of artificial intelligence provoked the most comment. The debate then moved smartly out of academia when the huge scale of Japanese investment in fifth generation research became apparent. In the UK, the Alvey Report and its recommended £350 million for information technology research is a direct response to the perceived Japanese challenge. Likewise, the books by Feigenbaum and McCorduck and by Simons take

ADVERTISEMENT



Tilapiine fishes of the genera *Sarotherodon*, *Oreochromis* and *Danakilia*. By Ethelwynn Trewavas.

A definitive account of these economically important mouth-brooding fishes, written with the interests of both taxonomists and fish culturists very much in mind. 1983, about 600pp, 188 illustrations. 0 565 00878 1. Hard covers. If ordered before December 1st. £45.00. Afterwards £50.00.

Seaweeds of the British Isles. Volume 1 Rhodophyta Part 2A Cryptonemiales (*sensu stricto*), Palmariales, Rhodymeniales. By Linda M. Irvine.

This is the second in a series of books that will when completed cover all of the British and the majority of northern Atlantic seaweeds. It will provide a standard reference work for the marine algae of the north Atlantic and beyond. 1983, 120pp, 28 illustrations. 0 565 00871 4. Water resistant soft covers. £13.00.

A new look at the dinosaurs. By Alan Charig.

All the known facts and theories about the dinosaurs are explained for the non-specialist with the aid of superb restorations, maps, charts and photographs. A series of colour plates depicts the animals in their likely habitats. First published (in hard covers) in 1979. This edition with amendments 1983, 160pp, illustrations include 16 pages of colour. 0 565 00883 8. Paperback £3.95.

Fossils — the key to the past. By Richard Fortey.

A lively and authoritative account of how the past is reconstructed from fossils, from microscopic coccoliths to the gigantic dinosaurs. The book is illustrated throughout in line, halftone and colour while the text includes notes on collecting, preserving and identifying fossil specimens. 1983, 176pp, illustrations include 32 pages of colour. 0 565 00884 6. Paperback £6.50.

Classification. By Susan Jones & Anne Gray.

A beginner's guide to three currently used systems of biological classification — phenetics, cladistics and orthodox classification. Each is clearly explained by drawings and diagrams to take the reader step by step through this complex subject. The work is based on the exhibit "Classification" at the Natural History Museum, South Kensington. 1983, 32pp, illustrated throughout. 0 565 00876 5. Paperback £0.95.

British Mesozoic fossils, 6th edition.

This latest up-dating of a classic work of reference incorporates extensive stratigraphical and nomenclatural revisions to bring it fully into line with current concepts. 1983, xi + 207pp, including 73 line plates, coloured map. 0 565 00872 2. Paperback £4.95.

Publications Sales, British Museum (Natural History), Cromwell Road, London SW7 5BD

their cue from Japan.

Geoff Simons' informative book is a helping hand for those perplexed by the thought of a fifth generation. He leads the reader so smoothly through the jargon that, although one does not emerge technically-speaking much the wiser, it does not seem to matter. This is no confidence trick — merely the effect of being led by an enthusiast who encourages long views. There is, in any case, a rich bibliography for the diligent. Whilst artificial intelligence and expert systems get a suitably prominent treatment, the book also covers the other likely components of future computer systems. Thus there is discussion of trends in integrated circuit technology, the rationale behind novel computer architectures such as dataflow, the use of programming languages such as PROLOG, and the possibilities opened up for man-machine communication by voice recognition and by visual perception. With Geoff Simons of the National Computing Centre as companion, one can enter the fifth generation debate and appear as intelligent as the next person.

The subtitle of the *The Fifth Generation* stamps it as a political statement. Professor Feigenbaum is one of California's Grand Old Men of artificial intelligence and Pamela McCorduck has long been an enthusiastic promoter of the cause. Their thesis, advanced with a touch of chauvinism, is that America invented the bulk of artificial intelligence technology and is now about to be overtaken as leader of the field. It is thus a plea for the US to spend more money on artificial intelligence research. To the computer scientist, this is a familiar cry which provokes adulation or derision, according to belief. To the innocent bystander, the arguments provide rare glimpses of a scientific sub-culture engaged in internecine strife.

Do not, therefore, read Feigenbaum and McCorduck to learn about fifth generation technology. Rather, read them for their perception of the evolving role of information processing in society, and their belief in the ultimate ascendancy of intelligent machines. If the background to their canvas is the whole of recorded history, the foreground is firmly occupied by what they state to be "the American predominance in everything from computing to finance, from industrial output to quality of life". This predominance is seen by Feigenbaum and McCorduck to be under attack from an unexpected source — Japan — and the message is that America ought to be worried. To a European audience, used to losing empires, the book's cries of righteous alarm could be mildly amusing if they did not at the same time contain a warning for all countries who aim to survive as industrial nations. The race to the fifth generation is on, and winner takes all. □

Simon Lavington is a Senior Lecturer in Computer Science at the University of Manchester.

Foibles programmed

Margaret Boden

In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension.

By Michael G. Dyer.

MIT Press: 1983. Pp. 458. \$40.50, £31.50.

If there is one prospect more alarming than having one's phone tapped by Boris, it is having one's diary read by BORIS. Denis Thatcher's resident Russian spy is a mere mortal, so may have a certain grudging sympathy for the sins that flesh is heir to. But BORIS's discretion in such matters cannot be counted on, for BORIS is a computer program. The potential threat to one's privacy arises because BORIS is a language-using program designed to understand narratives dealing with human plans, motivations, and emotions.

BORIS's data-base and inferential strategies represent some of our everyday knowledge about such matters as adultery and divorce, and the emotional (and legal) tangles that they may involve. Given the sentence "Paul wanted the divorce, but he didn't want to see Mary walk off with everything he had", this enables it to interpret "walk off with" as meaning possession rather than perambulation, and to see Paul's distaste for this prospect as a natural reaction to his discovery of his wife's infidelity.

Its representation of interpersonal phenomena such as anger, jealousy, and gratitude can lead it to see the point of an incident, without this having to be explicitly stated. For instance, it guesses the topic of a telephone call (Paul is asking his lawyer-friend Robert for advice) as it knows that present help may properly be requested and willingly granted because of a past favour.

The program also has some general knowledge of planning and of stereotyped behaviour: such as that clothes are usually kept in bedrooms, that to get to one's house from a restaurant one may need to be driven there, and that having the waitress spill soup on one's clothes may lead one to refuse to pay for the meal. All this heterogeneous knowledge is integrated in BORIS's processing (including its parsing) to enable it to give sensible answers to questions about a story in which a careless waitress leads to the discovery — with a witness — of a wife *in flagrante delicto*.

BORIS is the most impressive system based on the Schankian approach to natural language processing, for it integrates many different sources of knowledge in complex ways. A relatively novel concept is the "thematic abstraction unit", (TAU). TAU's are highly abstract schemata whose function is to organize memory and direct the process of understanding. They also support analogical reasoning and enable one story to remind

the system of another — superficially very different — one. Most TAU's are named by common adages, such as "a stitch in time saves nine", "too many cooks spoil the broth", and "many hands make light work". These are defined as patterns of planning and plan-adjustment which can be instantiated by many examples; associations within the program's semantic memory allow one example of "saving nine" to recall another.

The author's intent is to provide a psychological simulation of parsing and memory organization, and some experiments are cited accordingly. However, the psychological similarities may not go very deep. That stories are classified as "similar" with respect to their thematic nature (not their specific content) is a result which suggests that something broadly like TAU's are functioning within human

memory. But how broad is "broad"?

Only Boris, of course, could *really* understand one's diary: BORIS has no capacity for empathy or fellow-feeling. The program's "understanding" consists merely in manipulating symbols — including English words — so as to draw what human beings can interpret as sensible inferences about the interpersonal relations mentioned in the narrative. And only Boris could be bribed, with a suitable tincture, to use his discretion in passing on personally sensitive information. The teetotal BORIS's uncensored output might prove highly embarrassing if relayed to the Moscow end of a telephone wire. Perhaps Denis Thatcher should be thankful for small mercies. □

Margaret Boden is Professor of Philosophy and Psychology at the University of Sussex.

Point to point

C.J.S. Clarke

The Genesis and Evolution of Time: A Critique of Interpretation in Physics.

By J.T. Fraser.

Harvester Press/University of Massachusetts: 1983. Pp. 193. £19.50, \$20.

The Enigma of Time.

By P.T. Landsberg.

Adam Hilger, Bristol: 1982. Pp. 248. Hbk £13.95, \$28; pbk £7.95.

IS THERE such a thing as 'the riddle of time', whose solution would explain 'life, the universe and everything'? Or does there in fact exist a single phenomenon going under the name of time? It is the contention of J.T. Fraser that Time does constitute a unified subject of paramount importance, but that time is multiple and not simple.

His thesis is that the world of which we have knowledge can be thought of as organized into a hierarchy of levels, to each of which different laws of nature are appropriate, and in each of which time has a different character. Only at the human level does time have all the properties of a continuous past/present/future succession. The properties become progressively more rarified as one progresses to levels remote from our own, until at the lowest, most microscopic level, claims Fraser, there remains no time at all.

He realizes, of course, the need to show that this hypothesis can be made sufficiently precise for it to be tested against the actuality of nature. But unfortunately it is often hard to tell whether this approach really scores over more conventional ones: partly because he never enters into a detailed point-by-point critique of the alternative views; partly because the level of presentation is always very general, eschewing detailed mathematical analysis at every stage; and partly I suspect, because

he has chosen to ignore some aspects of the conventional physical view.

Consider, for example, the lowest, atemporal level (one of his most interesting ideas). As his starting point he takes the 'core' aspect of time to be metrical time, in a generalized sense. Time is that which is measured by a process of time-keeping, and time-keeping involves correlating events with some standard series of events chosen so as to yield a simple account of the phenomena. This is the tradition of time-explication started by Aristotle, beginning with the "numbering of movement" and then looking for a standardization of the numbering. It is sharply at variance with the approach of Newton-Smith, for example, who distinguishes topological time, a linear continuum without metrical properties, as the most basic level of structure for time. On this latter view one should separate the topology of time (involving questions of betweenness and of global topology) from the question of what metric one should place on a time with a given topology.

Fraser derives his lowest level from the consideration of null geodesics in special relativity. Such a geodesic has no intrinsic metric, but it has a perfectly good topology which coincides with the topological time of any inertial observer. And so it should make sense to see a photon as possessing its private topological time, though devoid of metrical time. For Fraser, however, time is metrical in essence and so the photon is timeless and inhabits, with any other massless particles, the atemporal level at the bottom of the hierarchy. The argument thus hinges on a philosophical issue as to the nature of time, where the alternative view (topological time) is not analysed at all.

The atemporal level is supposed to be the level in which the universe originated. From the point of view of cosmology this level lies at the temporal beginning of the universe, although strictly speaking it is without time and so cannot actually partake in a temporal relation. In placing the

atemporal level at the big bang, Fraser seems to take it for granted that in its earliest stages the universe was inhabited purely by photons, apparently ignoring the theories, now predominant, that place hadron physics (or quarks etc.) at the start.

On the central question of the directionality of time, the conventional dichotomy between laws (time-symmetric) and contingent conditions (imparting directionality) is rejected. Each level of nature has to be treated in its own right, and time is to be regarded as directional at that level if it is the case that the occurrence of a system obtained by reversing the time-sense of a normal system is not only unobserved, but is basically impossible. This, he claims, happens only at the biological level, where genetic programming and growth involve an unshakable directionality in time.

But what then is the relation, if any, between the basic time-asymmetry of the biological world and the contingent asymmetries of the physical world? It is acknowledged that the biological asymmetry has its ultimate root in the expansion of the universe, but the account of the link between the two is limited. For instance, although the paper of Penrose, arguing that the asymmetry is reflected in a progressive increase in the Weyl tensor which makes the Big Bang essentially different from the final Big Crunch, is cited, its conclusion is baldly denied by Fraser, who treats Bang and Crunch as equivalent.

Despite these drawbacks, the book has many attractive features. The metaphysics of his levels (his "umwelt principle") is a promising philosophical idea; the picture of an underlying atemporal world, notwithstanding its doubtful derivation, fits well with some speculative attempts to derive space-time from something more primitive in an attempt to avoid the non-renormalizability of quantum gravity; and the account of the origin of life, of primitive organisms as "clock shops" is intriguing. Fraser may well turn out to be right.

For an introduction to the conventional wisdom, particularly on the physical side, one could hardly better Landsberg's admirable anthology — *The Enigma of Time* — a selection of reprinted papers on irreversibility, cosmology, quantum theory and black holes with a concluding section on time in art and literature. Landsberg provides a lucid introduction, glossary and indices so as to make the book as accessible as possible to a wide readership. Inevitably, space prevents the detailed development of the themes; one's appetite is whetted for many more papers on the quantum measurement problem and its relation to irreversibility.

The anthology raises many questions, and Fraser's book attempts an answer. It is to be hoped that subsequent works will supply the critical comparisons that are needed to carry the debate further. □

Dr C.J.S. Clarke is a Lecturer in Mathematics in the University of York.

Penning a patchwork

Peter Kemp

The Whys of a Philosophical Scrivener.

By Martin Gardner.

William Morrow (Quill): 1983. Pp. 452.

Hbk \$22.50; pbk \$12.95.

MARTIN Gardner owns, he says, some fifty books each by G.K. Chesterton and H.G. Wells; he subscribes to *The Chestertonian Review* and "used to take . . . *The Wellsian*". Further testifying to his avowed "fondness" for these writers, his pages are packed with admiring references to, and admired references from, them. More basically, his entire enterprise in *The Whys of a Philosophical Scrivener* — writing as a jovial, voluble, at times almost self-parodyingly opinionated savant — seems considerably indebted to them. That Gardner should emulate authors who, he feels, "have long been out of fashion" is part of his determinedly unmodish persona. Ostentatiously at variance with anything that might be thought mercetriciously new-fangled, showing an almost cantankerous preference for those considered passé (William James is the philosopher most referred and deferred to), Gardner seems aptly published by Quill.

His quaint self-description as a "scrivener" presumably aims to match this robustly traditional image. A scrivener, though, was not an author but a copyist. And, in this sense, the word is more appropriate to Gardner's book than he may have intended. There is frequent reference to his "files"; and the book often gives the impression of someone dipping into dossiers of approvingly transcribed passages. It does not so much sustain a thread of argument as loosely tack together pieces of reasoning — or rhetoric — from those Gardner esteems: "I agree with William James that . . .", "Let us listen . . . to Unamuno", "here is Chesterton again".

This sometimes brings the book close to a patchwork anthology. And extracts, ripped out of context, can misleadingly represent their author. Gardner's treatment of Wells is a case in point. As a self-proclaimed theist, he deplores Wells's emphatic godlessness. Attempting to alleviate it, he refers with some frequency to the fact that Wells "was once so taken by the concept of a finite God that he wrote an entire book about it, *God the Invisible King*". What he doesn't make clear is that Wells harboured the notion only briefly, at a time of great nervous strain, during the First World War; that the God he envisaged was of a curiously belligerent cast — "as real as a bayonet thrust"; and that he was later shamefaced about his short period as a "theological Quisling".

Combining fairly recherché material — *God the Invisible King* is rarely perused

today — with an overriding theological drift, the technique is typical of Gardner. In his book, the "garnerings" of many years' wide reading bedeck an increasingly religious message: by the final chapter, readers are being urged to "Say something to God. Give thanks for something. Ask forgiveness for something." First sketching out his attitude towards truth, science, aesthetics, ethics, politics and economics, Gardner then devotes more than half as much space again to his feelings about God, prayer, evil, immortality. Mental gymnastics yield to the leap of faith. Mockery of current credulity about the paranormal gives way to spiritual affirmation: "I believe with my heart that God upholds all things".

Although possibly unexpected — "some readers", Gardner says, "may be surprised" at his theism — what emerges has a pattern. Expectably, in an admirer of Chesterton, Gardner relishes paradox. The unresolvable delights him. The epigraph to his book proclaims that "Philosophy is concerned with . . . soluble questions that are trivial, and crucial questions that are insoluble". And, as Gardner amply realizes, such crucial questions arise elsewhere, too. Quantum physics generates them. Some mathematical problems "can be 'solved' only by showing them to be unsolvable". In a way becoming popular, Gardner attempts to underpin theology with such insights. Applauding Gödel's incompleteness theorem, he remarks that it "can hit one like a religious conversion, bringing with it a great liberation from anxiety . . . I sometimes fancy that God invaded Gödel's mind (note the 'God' in his name)".

Initially, the book is characterized by a

taste for both reasoning and the reasonable. Fascinated by logical intricacies, Gardner subscribes to straightforward opinions; philosophical acuity is allied to blunt commonsense. A noticeably favoured way of reaching his conclusions is to steer between extremes. In politico-economics, he makes for the extensive ground between Adam Smith and Karl Marx. When it comes to mapping out his religious creed, a similar procedure is adopted. Investing his credo with an air of sensible moderation, Gardner locates it between the crankily paranormal and the orthodoxly religious. Both of these opposing extremes are ones he is well-placed to disparage knowledgeably. A keenness for assessing paranormal claims — from the Bermuda Triangle to "the power of the Great Pyramid's shape to preserve food and sharpen razor blades" — has given him rich instances of their poverty of conviction. Reared in an "ugly Protestant fundamentalism", he has a devastating knowledge of Biblical barbarity. Avoiding these outlandish extremes, his own dogma is, in many ways, as unexcitingly conventional as his political allegiances. Central to it is the familiar religious insistence on the limitations of the intellect. Religious faith, Gardner honestly emphasizes, is "belief unsupported by logic or science", a matter of emotional orientation: "emotional meanings play fundamental roles in decisions about philosophical questions". Behind Gardner's faith is hope. God, for him, "is essentially the provider of immortality".

As often, too, the nature of the deity revered reveals things about the worshipper. Characteristically, Gardner entertains speculations about a playful Provi-

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Autumn — an illustration from a new anthology, *Four Seasons*, which has been compiled by Edward Phelps and Geoffrey Summerfield (Oxford University Press).

dence. Disposed to the waggish himself — he attempts to trick the reader by coyly misattributed quotations; his favourite word is “whimsical”; he often refers delightedly to Frank L. Baum’s Oz stories — he would clearly welcome a deity whose disposition is similar. Perhaps, he hopefully suggests, the universe is “a vast cosmic jest fabricated by a god who had no motive except to amuse himself and his friends”. This cosy concept is discouraged, though, by memories from Gardner’s fundamentalist education: God only laughs four times in the Bible; on each occasion, he is chortling over the fate of the damned.

Moving from the Old Testament to the new physics, jumbling quarks and quirks, Gardner’s book is a weird blend of sophisticated cerebration and psychological naivety. In approach, it can be callow, full of sleeve-tuggings and leg-pullings. Being button-holed — “But wait! . . . Perhaps, dear reader . . .” — is a recurrent experience. Ratiocination roguishly

collapses into an exaggerated throwing up of hands: “Is it true? Don’t ask *me*. How could I possibly know? I put it forth whimsically”. As much a wise guy as a “whys” man, Gardner can be alienatingly facetious — yet this doesn’t stop his book from often being genuinely funny. Reverent gazings at the after-life co-exist with a sharp eye for the here and now. Political fatuities are savoured. And Gardner likes to pounce sardonically on Christian crassnesses — from the contemporary childishness of the born-again, carolling out that “Matthew 24 is Knocking at the Door” to medieval implements for uterine baptism of embryos. Oddly, a book introduced as simply “about what I believe and why” is at its keenest when debunking the bizarre beliefs of others. □

Peter Kemp is Senior Lecturer in English at the Middlesex Polytechnic, London. He is author of H.G. Wells and the Culminating Ape (Macmillan, 1982).

Who got where when

Nicholas Wade

Exploring the Earth and the Cosmos.

By Isaac Asimov.

Crown/Allen Lane: 1982.

Pp.352. \$13.95, £8.95.

ACCORDING to the dustflap of this book, Isaac Asimov’s two-hundredth book was published in 1979. The present work is doubtless his 210th, a fact that must make any reviewer lay aside, at least for a few seconds, the pose of omniscient scorn his kind is obliged to affect. A writer who has published 200 popular books is certainly a master of his craft, like it or not.

Exploring the Earth and the Cosmos is a craftsmanlike compendium of the history of geographical and extraterrestrial exploration. It is taut with pertinent facts, lucidly and deftly presented. It tells you the temperature of the planet Mercury, the distance to Alpha Centauri, the height of many great mountains and the depths of many profound abysses. It relates, with Guinness Book of Records exactitude, who got where first.

In this connection, Asimov may perhaps be a little too clement to the claims explorers make. He cites Admiral Peary as the first man to reach the North Pole, and since *Encyclopaedia Britannica* does likewise, he shall not be faulted for that. The astronomer Dennis Rawlins, however, showed in a book of 1973 that there is strong reason to doubt the good Admiral ever went near the pole. Since explorers tend to have few independent witnesses, but patrons and a public to satisfy, their claims often merit some scepticism.

It is here perhaps that Mr Asimov’s relentless march of facts begins to provoke

a twinge of fatigue in the reader. At some point people become more interesting than things to even the most hardened misanthrope. Someone had to be first to the North Pole; who cares if it was Admiral Peary or Eeyore? But if Admiral Peary was spinning a yarn, that’s a story that can be sat up for. Similarly the Venetian master Titian is earnestly cited for his longevity, but Titian was not the first to let his anticipation of a pension outweigh his appetite for factual exactitude. Modern scholars believe he was born in 1488, not 1477 as Titian and Asimov declare.

Expansion of physical horizons is the theme that threads the book together. Mr Asimov concludes by noting that “there are other horizons that have expanded — magnetic intensity, viscosity, angular momentum, etc — but those I have present are sufficient to show humanity at its most magnificently human . . .” The observation evokes the image of someone swimming through molasses in a 5 tesla field at 33g, which would surely be as fine a testament as any to the human side of humanity. The weakness of the book is that Mr Asimov’s interest in facts is inexhaustible, his attention to motive exiguous. Yet without motive, who would ever leave home to go exploring either Earth or cosmos?

Inclusion of the details that interest historians, however, would have extended the book beyond any reasonable scope. Mr Asimov’s aim was to distill the quintessence of fact, and he has succeeded excellently. Yet the book is probably one that more people will find useful as a work of reference than will read for pleasure. □

Nicholas Wade is an editorial writer on the New York Times and co-author of Betrayers of the Truth (Century, 1983), an analysis of fraud in science.

Smoothing the flow of data

T.A. Kletz

How to Write and Publish a Scientific Paper, 2nd edn.

By Robert A. Day.

ISI Press: 1983. Pp.180. Hbk \$17.95; pbk \$11.95.

MANY who have to read scientific papers, reports and memoranda, published and unpublished, will have sympathized with the chairman of a recent Inquiry, who wrote in his report, of the written evidence he had received, “Turgid and indigestible as some of it was, the Assessor and I read it, though often with little profit”. Many scientific and technical reports are badly written. (Is this why so few technologists seem to read books for pleasure? Does the garbage they have to read put them off all reading?) I therefore turned hopefully to the new edition of Day’s book. Could I recommend it to colleagues, past and present? It is well-written, it was a pleasure to read (I must get my tenses right), it contains many amusing anecdotes, but a better title would be ‘All you need to know about writing a scientific paper except what words to use’ or ‘How to make life easier for editors’.

This book covers choice of title, the abstract, the manuscript, the diagrams, layout of tables, references, proof-checking, ordering reprints and other such matters but only two short chapters deal with the actual choice of words. The author claims that if you organize the paper properly it will write itself, but this is not so. You could follow all his advice (except for the two chapters I have mentioned) and produce an unreadable paper; you could ignore all his advice and produce a paper that is a pleasure to read and which influences its readers, even though the tables take up more room than they need and the references are a mess. In fact, Day like many editors, almost forgets that it is the content that makes a paper and not the layout.

There is no reason why authors should not follow his advice and make his life easier — it is sloppy to have references arranged inconsistently, for example — but readers are (we hope) more numerous than editors and it is therefore more important to make their life easier by developing a clear style, particularly if you are writing for a wider public. Fellow workers in your field may persevere; others will not.

Day paints a picture of the editor as a man martyred by authors who will not follow his advice. As an author I do not quite see it that way; it is me (or my papers) that are afflicted. I could quote many examples of editors who have changed what they did not understand, shortened

articles so that they are finished at the bottom of a page and removed humorous examples that they considered too frivolous. I think of editors as dour, dark-suited Presbyterians, unwilling to let any humour into the serious business of imparting knowledge. I, on the other hand, agree with W. S. Gilbert that "he who'd make his fellow creatures wise, should always gild the philosophic pill", but my quotations from Lewis Carroll, my mathematical treatment of belt and braces to show how the reliability of two protective systems is calculated, have often been blue-pencilled.

Editors favour brevity; it is cheaper. But repeating a statement in different words often helps to get the point across. If the author of psalm 114 had been writing for a geological journal he would not have got away with "The mountains skipped like rams, and the little hills like lambs".

In emphasizing form rather than language Day almost treats scientific papers as art rather than a means of communication. Perhaps he is right to do so. Information scientists tell us that many papers are never read by anyone (except writer, editor and referee) and that at least 25% are never cited, even by the author. Just as medieval sculptors took great pains over the carving of cathedral roofbosses, that no one would ever see, so editors take great pains over the presentation of papers that no one will ever read. Depending on your point of view, both are a waste of time or art for art's sake: pride in a job done well, whether or not it is ever seen.

In a hard, commercial world is it really cost-effective to spend so much effort perfecting a paper to Day's satisfaction? Recently I spent many hours looking up the references to a longish review paper because the editor wanted more details than usual. It is sloppy, I agree, not to follow a journal's style, but how far do you go when life is limited? Do we aim for a perfect job or an adequate job?

The coming years may make much of what Day wants out-of-date. Already some publishers, as he points out, are setting type directly from the author's word-processor discs. This makes it harder for them to edit the text. The next step will be to cut out the typesetting altogether and feed the author's disc into a computer data base which readers can access, taking a hard copy if they want. Farewell, Mr Day.

In the meantime, his book should be widely available wherever scientific papers are written and all authors, particularly first-time ones, should be encouraged to read it. It will not solve all their problems; it will not solve their biggest problem, learning to write well, but it will smooth the path to publication and that is worth doing. Day has set himself a modest task and done it superbly. □

T.A. Kletz has retired from the chemical industry, and is now a professor in the Department of Chemical Engineering at Loughborough University.

Usage and abuse

Walter Gratzer

Fowler's Modern English Usage, 2nd edn.

Revised by Sir Ernest Gowers, 1965.

Oxford University Press: 1983.

Pbk with corrections. Pp. 725.

£3.25, \$8.95.

THE redoubtable Dr Burchfield is said to be embarking on a new edition of Fowler. Meanwhile Gowers' revision, available for the first time in sturdy paperback and a snip at the price, will do very well; no laboratory should be without it, if for no better reason than that a ramble through the scientific literature with Fowler as guide affords no end of macabre entertainment.

There is, as I understand, one scientific paper published every second. If Balzac could describe Georges Sand as a great cow full of ink, what are we to say about some of our *capi mafiosi*, who – I judge by some recent obituaries – are capable of putting out two thousand papers in a working life? (Who would have thought the old man had so much ink in him?). We live in a slop of overflowing neologisms, pleonasm, malapropisms, clichés and unattached participles.

Scientists, even more than journalists, batten onto each new distortion of language with an uncritical eagerness that they would never think to apply to the substance of their papers. Examples of current vogue words that come quickly to mind are 'probe', to mean any vehicle of observation, 'propose' to mean suggest, as in 'we propose that the core of the earth consists of molten Camembert', and 'abolish' to mean expunge, not a law or statute, but something on the lines of enzymic activity.

Even more striking is the way in which words with a precise technical meaning escape into the outside world, are mangled by journalists and politicians, and are then received back in their new and perverted sense. 'Parameter' for example is now widely used in the scientific literature, as in newspapers, to mean any measured quantity or constant (or, for that matter, variable), and 'extrapolate' and 'exponential' appear to be undergoing a similar malignant transformation. And how has 'multimer' suddenly surfaced, to signify apparently something that is smaller (or perhaps bigger) than an oligomer? (The proponents of this idiotic usage might not be aware of the old don's comment on television, when that medium began to assert itself: no good, he said, would come of an invention, the name of which was half in Latin and half in Greek).

Fowler is excellent on vogue words and admirably cool and dispassionate about such emotional issues as split infinitives and terminal prepositions. He concludes

his dissertation on the latter with the comment that not even Dryden, who denounced the practice, would have changed "which I will not put up with" to "up with which I will not put". (As a counter-example I would offer the child's question: 'what have you brought me that book to be read to out of for?').

Fowler is splendid on ambiguities. Here are two of his illustrations (presumably genuine) from the several categories that he considers: "to ask the Minister of Agriculture if he will require eggs to be stamped with the date on which they were laid by the farmer", and "Miss Pickhill grasped the pince-nez, which hung from a sort of button on her spare bosom". My own choices in this *genre* would be the exhortation on London Transport escalators that "dogs must be carried" (why should such an encumbrance be thought necessary?), the notice at one time displayed in the windows of a chain of chemists' shops, proclaiming "we dispense with accuracy", and the legend on the sides of removal vans, which used to be seen plying about London: "Removers of Distinction" it proudly asserted.

Fowler is a marvel of easy erudition and yields much diversion and instruction. I recommend it for browsing, for improving one's papers and resolving arguments and for confounding subeditors. I hope it finds its way across the Atlantic, whence the following, as an example of scientific prose to make your knees buckle: "Remarks as to the handiness of germane magnetic facts strongly parallel those made in antecedent paragraphs concerned with spectroscopic conduct. Generally speaking, magnetic demeanor is more greatly influenced by environmental vagaries than optical deportment, as the former is dependent upon the detailed visitatorial schedules of the transient electrons." Perhaps the computers are taking over. As that legendary master of the modern malapropism, Sam Goldwyn, nearly said, if Fowler were alive he'd be turning in his grave. □

Walter Gratzer is in the Medical Research Council Cell Biophysics Unit, King's College, London.

Gripes or wrath?

Our reviewer has given some examples of uncouth English usage, commonly found in the scientific literature. We propose to consider correspondence, in which readers can (briefly) air their prejudices, reflections and hates on scientific prose, its content, style and vocabulary. Contributions to our correspondence column on these matters are invited.

Clean lines

Ian Stewart

The Visual Display of Quantitative Information.

By Edward R. Tufte.

Graphics Press, Box 430, Cheshire, Connecticut 06410: 1983. Pp. 195. \$34 (US, postpaid).

EXPERIMENTAL scientists devote much of their time to obtaining data; theoretical scientists devote theirs to explaining data, or making predictions for experimentalists to test. A comparatively neglected stage is the presentation of data, which since the 1750s has increasingly meant graphical displays.

"Neglected" is perhaps not the right word: there must be few scientists who have not, at some time, drawn a graph or a statistical chart to illustrate their results or theories. But conscious attempts to analyse the processes involved in the effective communication of data are a rarity. This unusual and fascinating book is both a history and a critique of graphic display methods; the clarity of much scientific writing would benefit greatly if it were to become compulsory reading.

The first section of the book, "Graphical Practice", is a compilation of a wide range of examples of good graphic design. Edmund Halley's chart of trade winds and monsoons of 1686; the dot map made by Dr John Snow in 1854 which pinpointed the Broad Street water pump as the source of a cholera epidemic; an amazing time-series

plot of planetary movements dating from the tenth or eleventh century; Marey's train schedule for the Paris-Lyon run in the 1880s; and Minard's *tour de force*, a chart of what befell Napoleon's invasion of Russia, which displays six variables simultaneously. A pivotal role is assigned to William Playfair, who assiduously sought to replace tabulations of statistical data by graphs and charts, and was responsible for many innovations of graphic design that are now taken for granted.

The author lays down several guiding principles for graphical excellence, of which two will suffice here: "Graphical excellence consists of complex ideas communicated with clarity, precision, and efficiency"; and "Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space".

There is the (almost obligatory) analysis of how graphics can be used to distort statistical data, a crime committed by Playfair himself as part of a polemic against the supposedly skyrocketing national debt. In particular, he failed to allow for inflation — a common sin even in these enlightened times. This is familiar material, but it is well documented and bears careful reading.

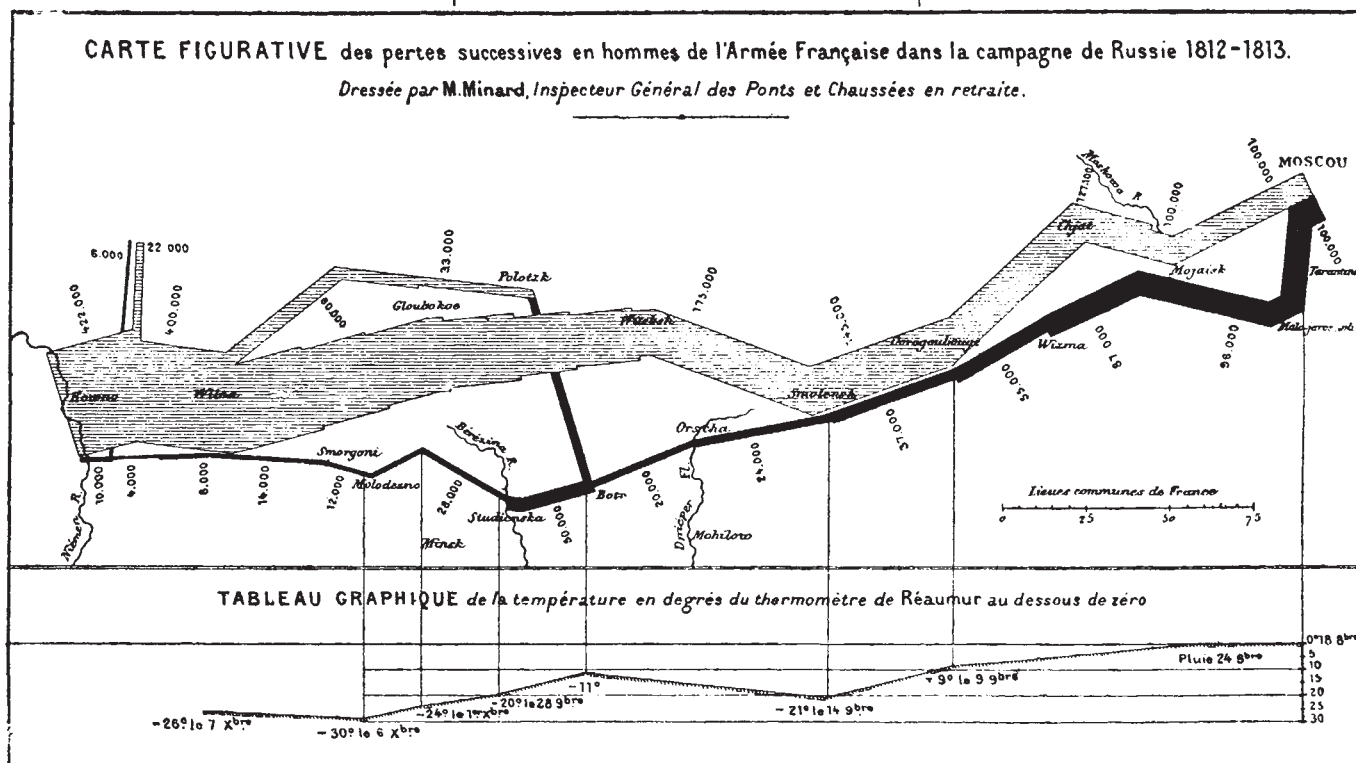
The heart of the book, though, is its second, more critical part, with the somewhat off-target title "Theory of Data Graphics". In reality it is a devastating critique of many standard graphical techniques; but a constructive one that suggests many ingenious and effective improvements and alternatives. However, many of these are dependent on very recent advances in printing technology.

There is a scathing and accurate attack on the "We Used a Computer to Build a Duck" syndrome — a duck being the author's symbol for a graphic that subjugates content to style (especially in a pretentious and "arty" fashion). The name is taken from the Big Duck Store in New York, shaped like the bird in question, for which "the whole structure is itself decoration". Today's computers make the production of immensely complicated multicolour multidimensional charts relatively routine: this is a timely warning indeed, deserving more attention than (sadly) it is likely to get.

The author takes his dictum "with the least ink" to heart. Many arguments against particular designs are supported by an analysis of the percentage of ink devoted to "chartjunk" — parts of the display that convey no information (or, worse, obscure the parts that do, such as unduly heavy grid-lines, eye-boggling Moiré-effect shading, or irrelevant frames).

It is not necessary to agree totally with the author's re-designs to accept that many standard displays (such as the awful pie-chart) are cluttered, fussy and fail in their aim: to convey information more clearly than a bare table of numbers, but as honestly. There is much here that will sensitize readers to the problems and reveal defects in designs; and as scientific data become increasingly multidimensional, effective and honest simplification must become the order of the day. □

Ian Stewart is a Lecturer at the Mathematics Institute, University of Warwick.



The terrible fate of Napoleon's army during the Russian campaign of 1812. The band widths show the size of the army over time and space. The hatched band on the left shows the start of the campaign, at the Polish-Russian border. The dark band shows the retreat from Moscow. An early example of displaying multiple variables.

Standing up for science

Eric Ashby

A Science Policy for Britain.

By Tam Dalyell.

Longman: 1983. Pp. 131. £5.95.

Two years ago a Select Committee of the House of Lords in Britain published a report on *Science and Government*. Among the witnesses to the Committee was the President of the Royal Society of London. He was asked how often a scientist from a government department came to the Royal Society to ask for opinions or advice or factual information. "Is it so frequent?" he was asked, "as to be a daily occurrence or so infrequent as to be a memorable occurrence?" The President's reply was: "It is perhaps nearer the latter than the former". Representatives from British universities were asked a similar question and they gave a similar response: "...there is [in universities] this undoubted reservoir of what we regard as relevant talent which is largely untapped."

The so-called "science vote", i.e. the amount of public money allocated by the government for the support of science, amounts to nearly £600 millions a year. No surprise, therefore, that politicians and civil servants get spasms of unease about the efficacy of science in the service of the nation. We are witnessing such a spasm now. *Nature* (304, 383) has reported recently the findings of a committee set up by the Advisory Board of the Research Councils to examine the deployment of their segment of the science vote; and, a week later, (*Nature*, 304, 476), it is announced that the government has commissioned Sir Ronald Mason to carry out an inquiry into the organization of the

research councils themselves.

The last spasm of unease was treated by Lord Rothschild twelve years ago. In vigorous and forthright prose he prescribed a cathartic to wash out the complexities of science policy. 'It's simple', he seemed to say. 'Treat scientists as you treat plumbers. The customer wants a job done. He pays a contractor to do the job: a principle as valid for the control of water pollution in the River Thames as it is for the control of the water closet in your bathroom.' It turned out not to be as simple as that. The 'customers' often do not know what to order; the 'contractors' have to waste time on bureaucratic chores and have to accept priorities which interrupt the continuity of research upon which the nation's future prosperity may depend. The Royal Society's verdict on the present procedures (given in its evidence to the House of Lords Select Committee) is unambiguous: "present arrangements for interdepartmental co-ordination of scientific advice to Government are much less satisfactory than those which were in operation in 1971".

So science policy is going to be featured in the news in coming months. Tam Dalyell's little tract is a useful guide to the issues that need to be discussed: the labyrinthine pathways in the Whitehall Village (Whitehall being a term-of-art for the British Civil Service); the purpose and influence of royal commissions (Michael Foot said they were "like a broody hen sitting on china eggs"); the machinery for distributing and using the science vote; the contribution made by British science to the Third World; the frustrating gap between a promising technical discovery and its exploitation by industry. Dalyell is better informed than most members of parliament about the state of science in Britain. He acquires information avidly; for years he has distilled his impressions in an admirable highbrow gossip column in the *New Scientist*. And, in contrast to scientists

who dabble in politics, Dalyell (who would willingly concede that he dabbles in science) really does understand politics.

The reader should approach *A Science Policy for Britain* as though he was going to meet Tam Dalyell for conversation over a drink. His book is not well organized or tightly argued. It is mildly party-political: material that might be considered in some future manifesto for the Labour Party. It has two welcome merits: it is clearly not a mere compilation from books and reports; Dalyell has talked to workers at the laboratory bench and he

writes from personal experience. And — as you would expect from such an experienced politician — it does not offer Utopian recipes of the kind that any civil servant would put into the waste paper basket.

Indeed, Dalyell's recipes are inevitably the most disappointing pages in the book: "To ensure that within the Science and Engineering Research Council that research in science and engineering forms a continuum"; "To find ways of improving the lot of universities and colleges. . ."; "To establish a firm framework for telecommunications research in Britain". These bland and sedative sentiments might well have come from some timid committee but they are out of character from the pen of such an articulate propagandist as Tam Dalyell. The title of his book would have been more descriptive of the content if it had been: "The need for a science policy in Britain".

The value of the book, apart from the facts in it, is (for me) the implicit disclosure of the politician's dilemma. He knows what is wrong. He has good ideas about how to make things better. But he is up against his own colleagues: "Sadly my experience of British politics leads me to the sombre conclusion that what is immediate takes precedence over and above what is important". And he is up against the impregnable defences of the civil service: its reserves of accumulated wisdom and know-how; its flair for tactical procrastination; and its obsession with confidentiality. There is an ironic reminder of this in the volume of evidence given to the House of Lords Select Committee in the report I have referred to. Look down the page numbers on the table of contents. When you come to the evidence from the secretary of the cabinet and the head of the central policy review staff you read the words "not printed".

In evidence to that same Select Committee Sir Alan Cottrell identified the basic problem that has to be solved before there can be a consistent policy for science in Britain. There are, he said, two models for governments to adopt. One, adopted by the Germans and Americans, is to "create broad conditions under which industry, by its own efforts can prosper". The other model, adopted by the French and Japanese, is much more dirigiste toward industry. The latter system — towards which British governments have in the past been inclined — demands a civil service totally different from the one we have at present. Until this issue is clear it shall not be able to design a science policy for Britain. But let us be grateful to Tam Dalyell for briefing us about the issues in this book. □

THE MINISTER'S JOB IS TO
BLIND POLITICIANS WITH
SCIENCE AND SCIENTISTS
WITH POLITICS!



Lord Ashby was for two years a member of the House of Lords Select Committee on Science and Technology.

Phrenology revived

John C. Marshall

The Modularity of Mind.

By Jerry A. Fodor.

MIT Press/Bradford Books: 1983.

Pp. 145. Hbk \$17.50, £15.75;

pbk \$8.50, £7.65.

IN THE popular imagination Franz Joseph Gall (1758–1828) is usually assigned a prominent position in cranks' corner. And indeed a man who could purportedly diagnose talent and character by feeling the bumps on one's skull does not seem too well placed to assume the status of scientific hero.

For the most part, Gall's life and work has supported two industries into whose clutches only a fiend could despatch his worst enemy: first, Gall has been grist to the mill of those philosophers who are convinced that they can infallibly distinguish science from pseudo-science; second, he has provided a happy hunting ground for sociologists who can see parallels between the cultural role that phrenology played in the Victorian era and the recent rise of Californian freak religions. Such are Gall's detractors . . . but his current supporters are typically an even more peculiar lot. The last book I read on phrenology was written by a lady who described herself as "a practising witch".

What then could provoke an eminent Professor of Psychology and Philosophy at MIT to start (albeit metaphorically) "hanging around with phrenologists and other dubious types"? The first thing to make clear is that Jerry Fodor has *not* taken to the laying on of hands: everyone agrees that craniology was a mistake, although an understandable one in the days before CT scans, positron emission tomography and nuclear magnetic imaging permitted *in vivo* visualization of the human brain. Rather, Fodor's primary concern is with Gall's views on *how* the complexity of mental life (and its neuronal substrate) is governed by the operation of a (relatively) small number of discrete components. To appreciate the force and originality of those views we must first look at the tradition Gall came to subvert.

In order to understand the human body we divide it into organs and organ-systems, and into tissues of different types, each with distinct functions. To paraphrase Thomas Hobbes: what is the heart but a pump, and the kidneys but filters; the muscles contractile springs, and the peripheral nerves but so many irritable strings? The basic mechanisms are isolated and the whole interpreted as the interaction

of its parts. For at least two millennia, students of cognition have essayed the same strategy: to analyse intellectual capacity as the interplay of diverse primary powers. Methodologically, the paradigm is impeccable. How else could one possibly proceed? But substantively, the utility of the strategy clearly presupposes that we have correctly carved nature "at its joints".

The traditional fractionation of the mind derives from Aristotle and the Early Church Fathers, for whom the faculties of memory, reasoning, attention, perception and abstraction are the fixed inborn components that underly mental activity. These faculties are individuated with respect to their effects, not their subject matter. Thus the principles of memory are,

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

on this account, indiscriminately applicable to the storage and recall of events, propositions, faces or melodies; in like fashion, one learns to recognize both roses and Mozart symphonies by abstracting the common elements from the class of exemplars to which one has been exposed; one solves problems in mechanics and the dating of prehistoric artifacts by application of the self-same laws of reasoning to each different data-base.

Contrary to vulgar myth, no exponent of faculty theory ever believed that merely naming the powers of the mind constituted an explanation of those powers; Molière's little joke that opium causes sleep because it possesses "dormative virtue" was singularly misguided. The point of the enterprise was to delimit the activities for which future research might hope to discover the *specific* nature of the mechanisms responsible for those activities, their quantitative parameters and computational structure. Nor should one imagine that the theory is of merely

historical interest. Whenever a modern theoretician postulates a memory system whose capacity is limited to n items, irrespective of what those items are, there lurks a traditional faculty; whenever a perceptual theorist argues that objects *in general* are recognized by computing the "distance" (in some innate quality space) between the stimulus presented and the members of a set of stored "prototypes", there we have an Aristotelian faculty. In short, Fodor points out, the building blocks of classical psychology *and* much contemporary psychology are mechanisms whose *modus operandi* is constant over whatever content is input to them.

And so to Gall. Gall argued, point blank, that the faculties advanced by all previous theoreticians were, in all probability, a fiction. For Gall, there is, as Fodor writes, "no such thing as judgment, no such thing as attention, no such thing as volition, no such thing as memory". A glance at any phrenological bust will confirm that there is literally no overlap with the old faculties. With respect to higher cognitive functions, Gall's innate mental organs are individuated solely in terms of the specific content domain with which they deal. What we find are organs of calculation (= mental arithmetic), of locality (= the representation and recollection of space and place), of tune (= musical harmony and melody) and of language. Gall did not, of course, deny that one could perceive a melody, attend to a calculation or recall a word. The phrenological *credo*, repeated in every text, was simply: "That attention, perception, memory, and imagination, are not primitive faculties of the mind, but only

modes of activity of all or any of the intellectual faculties". It then follows from Gall's commitment to a punctate locus within the brain for each *real* faculty that the frontal lobes cannot be the seat of reason, nor the temporal lobes the seat of memory. And hence the fact that organs of tune, locality and language *et cetera*, are so carefully delineated on phrenological heads. Gall's position also implies that it is entirely an empirical issue whether the mechanisms involved in the perceptual analysis of sentences have anything in common with the mechanisms implicated in the perceptual analysis of faces.

So the fundamental powers that Gall conjectures are domain-specific, genetically determined, associated with discrete neuronal structures and computationally autonomous in that all the machinery necessary for their operation is contained within each distinct mental organ. Fodor coins the term "vertical faculty" to refer to Gall's organs, which leaves him with "horizontal faculty" for

the traditional concept of domain-independent mechanisms wherein different "contents" compete for limited processing resources. "It seems to me", Fodor writes, "that the notion of a vertical faculty is among the great historical contributions to the development of theoretical psychology", an evaluation with which I fully concur.

A more pointed issue, however, is: are Gall's views correct? Fodor now advances a taxonomy of mentation that specifies those aspects of mind that do seem to be "modular" (in Gall's sense) and those that do not. Fodor distinguishes between transducers (the sense organs that encode proximal stimulation into neural signals), interface systems that compute the structure and character of the distal objects that correspond with particular proximal stimulations, and central systems concerned with the fixation of belief. Interface systems, Fodor argues, are modular, central processes are not. The interface systems are the various mechanisms of domain-specific perceptual analysis *plus language*, a somewhat "odd category", as Fodor remarks, from the viewpoint of any traditional fractionation of the mind. What does unify the class functionally is that its members all "represent the world as to make it accessible to thought", although they do so in radically distinct ways: "What underwrites the correlation between visual stimulations and distal layouts are (roughly) the laws of light reflectance. Whereas, what underwrites the correlation between token utterances and distal layouts is (roughly) a convention of truth telling". Granted that Fodor's functional characterization is correct, we can then inquire why, in the evolution of the brain, modular systems should have been selected for precisely this class.

The operation of systems that "present the world to thought" must be fast and mandatory if the organism is to survive in a world of predators, moving cars or academic discussion. Interface systems achieve speed by having built in to them a "theory" of the domain to which they are responsive. The algorithm that derives the intrinsic shape of a moving object is constrained by the principle of rigidity: if a collection of moving points has a unique interpretation as a rigid body in motion, "see" that interpretation. Similarly, effective algorithms for parsing sentences must have available to them the structural principles of universal grammar. (Chomsky's Cartesian notion of innately known principles thus fits well with Gall's notion of a "vertical faculty".) The domain-specificity of cognitive modules therefore follows from the specificity of the types of information they must contain.

Fodor further argues that interface systems are "informationally encapsulated" in that their mode of operation does not have access to high-level information about what is likely to be seen or said: "a condition for the reliability of perception,

at least for a fallible organism, is that it generally sees what's there, not what it wants or expects to be there. Organisms that don't do so become deceased". Fodor's discussion of informational encapsulation draws mainly upon the perception and comprehension of language, where the dominant view (in both artificial intelligence and experimental psychology) is that all kinds of semantic and pragmatic information are brought directly to bear upon the acoustic signal. Fodor's careful unravelling of what exactly the relevant experiments do and don't show should ensure that this free-for-all approach is quietly dropped.

In sum, Fodor defends and extends a Gallist approach to the architecture of

cognition that is more detailed, more tightly argued and better supported by experimental evidence than Gall himself could have dreamed of. Even at the height of his fame as an anatomist, no respectable scientist would overtly admit the validity of Gall's approach to the fractionation of the mind. Jerry Fodor has never been noticeably respectable, thank goodness; he writes too amusingly for that. But the thesis Fodor advances is the current monograph may well be right. And that, plus having written the most lucid and incisive analysis of theoretical psychology thus far conceived, should be consolation enough. □

John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at the Radcliffe Infirmary, Oxford.

Use and abuse of racial differences

W.F. Bynum

Race and Racism.

By Ruth Benedict.

Routledge & Kegan Paul: 1983. Pp. 192.

Pbk £3.95.

OF only one science could it be proposed that its two foremost American practitioners in the middle third of the present century were women. The science is anthropology and the women were Ruth Benedict (1887–1948) and her pupil Margaret Mead (1901–1978). Of the two, Mead undoubtedly had the larger measure of popular success, ultimately becoming a kind of media guru, venerated for her wisdom derived from the study of other societies but called into service in understanding her own. By contrast, Ruth Benedict was of a retiring disposition, and in any case, the fact that she was fourteen years older than her pupil meant that her own academic career was pursued at a time when women rarely occupied chairs in major universities.

Benedict was promoted to a full professorship at Columbia University only shortly before she died. Yet, her book *Patterns of Culture* (Houghton Mifflin, 1934) has been described as the single most influential work by a twentieth century American anthropologist. It has remained in print and been translated into fourteen languages. In this book, in measured prose and with remarkable insight, she used a series of case studies of American Indian cultures to develop a subtle thesis about the ways in which societies mould

the individuals born into them. She pointed out that not just the values and mores, but the perceptions and behavioural strategies can be radically divergent in different cultures: and in her particular case studies, in people who, biologically speaking, belong to the same 'race'.

By the time Benedict published her *magnum opus*, the Nazis had already come to power, though her book made no reference to the wider world of European politics. Its message, however, was consistent with that of *Race and Racism*, reprinted now after forty years. Originally published in a somewhat fuller version in America in 1940 as *Race: Science and Politics*, this book was explicitly written for a world drawn into a war in which racism played a central role. Benedict's aims were clear: to explode the myth that race and culture ever have been or ever can be congruent entities, and to argue, using biology, psychology and history, that race can never be used as a valid basis of categorizing individuals.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Prefiguring death — a sketch of Ruth Benedict in 1948 by Erick H. Erikson (from An Anthropologist at Work by Margaret Mead, Houghton Mifflin).

Race, Benedict insisted, is a biological fact; it can be studied using ordinary laboratory techniques. One's 'race' is genetically determined at conception, just like one's sex. The investigation of race and racial differences is a laudable enterprise, for it can teach human beings more about themselves and Benedict was sanguine enough to believe that self-knowledge was a salutary goal. Despite the biological reality of race, however, she pointed out that scientific efforts to divide the human species into particular races could never yield any universally satisfactory result, because of what she called the 'mingling of peoples'. The nature of history was such that 'pure' races do not actually exist, and the larger the racial grouping (such as tripartite division into Caucasian, Mongoloid and Negro), the fewer the physical characteristics which can be more or less consistently accommodated. Further subdivision permits the taxonomic inclusion of additional characters, but at the price of an increasing number of anomalies.

Two particularly cogent themes emerged from Benedict's reflections on the nature of race and racism. First, she was firm that there was no evidence of a difference in intelligence between the various races. Variation in performance on intelligence tests she attributed entirely to cultural or social factors. For instance, some cultures hold that it is impolite to answer a question unless one is absolutely and unassailably certain. Members of these groups will uniformly score badly on conventional intelligence tests. In other cases, scores improve dramatically when the children of a race (e.g. Black Americans) are offered the social and educational opportunities of the Northern United States, instead of the South. She did not discount the possibility that there are significant genetic components to intellectual capacity, but individual variation was so great and human beings so culturally pliable that generalizations based on race were meaningless. She based her statements on research done in the 1930s, but modern readers will find her discussion topical.

A second theme addressed the issues surrounding the rise of Nazism and the 'Aryan myth'. Her point, argued primarily from history, was that racism (the notion that one group has the stigmata of superiority and the other has those of inferiority) was a relatively modern phenomenon without biological foundation. It was, as she put it, the 'new Calvinism'.

Although Benedict wrote as a scientist, her inclinations, one senses, were really

those of a humanist. (It could be argued that the two designations are not incompatible.) She had a passionate commitment to human dignity and equality, and four decades of debate and research have done nothing to detract from the essential core of her analysis.

A foreword by Professor John Rex puts this little book into contemporary pers-

pective, and two statements on the biology of race prepared in 1964 and 1967 by scientists convened by UNESCO round out the volume, which retains interest both as a period piece and as a clear exposition of its subject. □

W.F. Bynum is at the Wellcome Institute for History of Medicine, London.

Neural counterpoint

Steve Blinkhorn

The Throwing Madonna: Essays on the Brain.

By William H. Calvin.

McGraw-Hill: 1983. Pp.253. Hbk \$13.95; pbk \$7.95.

SOME see popular science as simply the dissemination, in a stripped-down form, of what one might call regular science. 'Sexy' science has a simplified outline. There are many styles in which this is done, from Scientific American didactic, through Cousteau's heavily personalized 'denizens of the deep' approach to Reader's Digest 'I am John's lateral geniculate nucleus'. On the other hand there is a genre which seeks to dance a light-footed counterpoint to regular science, to act not merely as a means of transmission but to add to and cast into new perspectives what lies unloved by all but a few in the more impenetrable journals.

The Throwing Madonna is a case in point. If one can get past the title and stomach the overly self-conscious, occasionally arch and always unlovely prose style, one finds a thesis advanced, a point of view explored and a deal of loosely related material drawn together by the overall theme of 'Essays on the Brain'.

The thesis, that handedness, stone-throwing, the development and control of language and the practice of mothers carrying infants on the left arm, are linked in the origins of our species belongs to what one might call speculative evolutionary paleoanthropology. This is a field refractory to attempts at empirical disconfirmation, and rich in connotation. It is the world of the artist's impression, in which hominids more hirsute than ourselves are busy inventing tools in the foreground, whilst in the background still more hirsute apes look on uncomprehending, but somehow sullenly envious. Calvin's peculiar merit is that he pursues his thesis into the structure and function of the cerebral cortex of modern man, in a way which holds out the distant prospect of a solid evidential basis for his ideas.

Once into the modern brain he is considerably more at home, examining the relevance or otherwise of the digital computer as a metaphor for the brain,

dislodging the action potential from its preeminent place in the process, and ranging over tic douloureux, schizophrenia and the inadequacy of funding for research in his subject. But all too often one finishes an essay with a sense that more might (and should) have been said, and the considerable unevenness of style suggests that if the book of essays has indeed reached maturity as a form, then Calvin is not its most practised exponent.

There is a point of view that works such as this are essentially disreputable, a kind of intellectual self-indulgence which has no place in the rigorous world of science. The contrary view, that it is in the less disciplined exposition of ideas which ultimately drive more rigorous research that the sense and direction of science are found, receives some support from this at times engaging collection. It is a modest product, worth its modest price. □

Steve Blinkhorn is Director of the Psychometric Research Unit, Hatfield Polytechnic, UK.

ADVERTISEMENT

Published on the
centennial of the eruption

Krakatau 1883

The Volcanic Eruption and Its Effects

TOM SIMKIN AND RICHARD S. FISKE

History's most famous eruption, with:

- Ashfall over 6000 km away
- Sea waves over 30 m high
- Explosions heard 4653 km away
- Dramatic sunsets around the globe
- Return of life to island remnants

These events are described in the words of 88 eyewitnesses and interpreted in 28 selected scientific accounts written over the last 100 years.

464 pp., 139 figs., 49 full color illus.
Cloth: ISBN 0-87474-842-9 \$25.00
Paper: ISBN 0-87474-841-0 \$15.00

Order from:



Smithsonian Institution Press
P.O. Box 1579
Washington, D.C. 20013

International orders accepted. All orders must be prepaid by check, money order, or credit card in U.S. dollars only.

Circle No.09 on Reader Service Card.

The next review supplement to appear in *Nature* will be Christmas Books on December 1.

Books to be reviewed include *Computer Images: The State of the Art* by Joseph Deken, *The Stars and the Stones* by Martin Brennan, *Landscapes of the Night: How and Why we Dream* by Christopher Evans and *Fossils: the Key to the Past* by Richard Fortey.

Bones, teeth and molecular clocks

Bernard Wood

New Interpretations of Ape and Human Ancestry.

Edited by Russell L. Ciochon and Robert S. Corruccini.

Plenum: 1983. Pp. 888. \$95, £66.50.

PALAEoANTHROPOLOGICAL research over the last decade or so has been rather like an iceberg. Above the metaphorical waterline has been the veritable flurry of new discoveries and fresh interpretations of hominid remains from the Plio-Pleistocene, but beneath the surface, literally and metaphorically, lies the Miocene, which, in the same period, has been the focus of much fossil collecting, field research and careful laboratory study and analysis. The editors of this weighty book quite rightly recognized the need for a summary of the progress and the problems associated with research into this crucial phase of hominoid evolution and this volume is the result. Any doubts about its adequacy can be immediately dispelled. The standard is high and the editors have obviously worked hard to produce a volume which is every bit as good as Washburn's *Classification and Human Evolution* (Aldine) which, in 1963, covered similar, but obviously not identical, ground. Although the majority of the papers in this volume (some 30 in all) concentrate on fossil evidence and its interpretation, the scope of the book is wider, and its real theme is any evidence, be it molecular, chromosomal, palaeontological or palaeoenvironmental, pertinent to an understanding of the emergence of higher primates.

At the beginning of the 1970s, the majority of palaeontologists considered that the lineages leading to man, and the African and Asian apes, had been

established by the middle of the Miocene epoch (25–5 million years BP), if not before. *Ramapithecus* was a hominid ancestor, and all was well with the world. However, the middle and late 1960s saw the beginning of the collaboration between Vincent Sarich and Allan Wilson at Berkeley. Starting with immunological techniques, they embarked on a research programme which aimed to assess the strength of what can best be termed the biochemical or molecular relationships between extant primates. Sarich was convinced from the onset that "the essential variable governing the retention of antigenic similarity and appearance of antigenic diversity is time". They developed the notion of the 'molecular clock' and the time it told for the divergence of the higher primates was not 20 or 15, but 5 million years. What progress at reconciling these two scenarios, if any, has been made in the past decade or so?

The practitioners of molecular evolution would claim, I suspect, that it has taken the palaeontologists that long to see the light and that they are merely using new fossil evidence and the fruits of more rigorous morphological analysis as a metaphorical smokescreen to cover their unseemly and ragged retreat. This view is both unfair and untrue. Palaeontologists were indeed defensive of their position, but they were also rightly sceptical of the more extreme claims, or alleged claims, which were being made for the molecular evidence.

Although molecular evidence was closer to the genome, it was nonetheless still phenotypic evidence. However, some workers treated it as something akin to absolute truth, whereas it was clearly no such thing. Molecular evolution is a branch of comparative anatomy — admittedly the anatomy of molecules, but nonetheless still susceptible to the same class of errors as more traditional comparative studies. However, whatever ones caveats about the molecular evidence, its advent most certainly made palaeontologists examine the fossil record more closely than they

might otherwise have done, but more importantly it prompted them to embrace more rigorous methods of analysis.

What followed, in the field of molecular evolution, but particularly in palaeontology, is well represented in this book. The pre-Miocene evidence from the Fayum is reassessed in one chapter and it is claimed that *Aegyptopithecus* and *Propliopithecus* are suitable "phyletic ancestors for all later catarrhines". The hominoidea are regarded as the primitive catarrhines from which the cercopithecoids evolved as a derived group. *Ramapithecus* and *Sivapithecus* are treated by many contributors as synonymous, and linked phylogenically with modern and sub-fossil *Pongo*. Two authors conduct a spirited rearguard action on behalf of the hominid affinities of *Ramapithecus* and there is now clearly a need for a detailed examination of the evidence cited by Kay and Simons in the light of the well documented counter-evidence put forward in this volume and elsewhere by Andrews, Tekkaya, Ward and Pilbeam. Wolpoff contributes a thoughtful review in which he postulates that the *Ramapithecus/Sivapithecus* 'radiation' may include the common ancestors of all the living hominoid forms. Contributions by Mai on chromosomes, Kluge on cladistic analysis and Kortlandt on the palaeoenvironmental evidence were papers which particularly caught my interest. Some authors are guilty of using cladistics more as a prop than a tool, but the majority were rigorous and comprehensive in their treatment of their elected or allotted topics.

I have a few 'niggles' for the editors, but many more plaudits. I would have appreciated more maps and timecharts at the front or the back of the book and a summary of each paper would have been helpful in a book of nearly 900 pages. This must be one of the first books on hominoid palaeontology in which new words and new meanings, outnumber new taxa. Diagram is now a verb, so I can now spend my time 'diagramming' if I tire of writing. 'Meld' is a new and subtle combination of weld and mould, and if I get bored with 'diagramming', I can always spend my time 'reconceptualizing'! More seriously, though, the editors can be heartily congratulated on their efforts. Mistakes are few and minor. There are comprehensive indexes to both authors and subjects and one of the editors has provided an excellent concluding summary chapter.

This book is edited by two students who graduated from the Department of Anthropology at Berkeley. They dedicate it to two of their teachers, Sherry Washburn and Clark Howell whose ideas and influence permeate much of what is good in palaeoanthropology today. I would urge anyone seriously interested in human origins to read this excellent book. □



The orang-utan — a close but disputed relative of human beings. The picture is taken from *A Complete Guide to Monkeys, Apes and other Primates* by Michael Kavanah, (Johnathan Cape), which will be reviewed in *Nature*.

Bernard Wood is Professor of Anatomy at The Middlesex Hospital Medical School, University of London.

Squeezed states of light

D. F. Walls

Physics Department, University of Waikato, Hamilton, New Zealand

The properties of a unique set of quantum states of the electromagnetic field are reviewed. These 'squeezed states' have less uncertainty in one quadrature than a coherent state. Proposed schemes for the generation and detection of squeezed states as well as potential applications are discussed.

THE electric field for a nearly monochromatic plane wave may be decomposed into two quadrature components with time dependence $\cos \omega t$ and $\sin \omega t$ respectively. In a coherent state, the closest quantum counterpart to a classical field, the fluctuations in the two quadratures are equal and minimize the uncertainty product given by Heisenberg's uncertainty relation. The quantum fluctuations in a coherent state are equal to the zero-point fluctuations and are randomly distributed in phase. These zero-point fluctuations represent the standard quantum limit to the reduction of noise in a signal. Even an ideal laser operating in a pure coherent state would still possess quantum noise due to zero-point fluctuations.

Other minimum uncertainty states are possible which have less fluctuations in one quadrature phase than a coherent state at the expense of increased fluctuations in the other quadrature phase. Such states, which have been called squeezed states¹⁻⁵ (other names include two photon coherent states, generalized coherent states), no longer have their quantum noise randomly distributed in phase. Such states offer intriguing possibilities. In the present optical communication systems which use coherent beams of laser light propagating in optical fibres, the ultimate limit to the noise is given by the quantum noise or zero-point fluctuations. If, instead, beams of squeezed light were used to transmit information in the quadrature phase that had reduced fluctuations the quantum noise level could be reduced below the zero-point fluctuations. Optical communication systems based on light signals with phase sensitive quantum noise have been proposed by Yuen and Shapiro^{6,7}.

The concept of squeezed states applies to other quantum mechanical systems. For example, they may have a role in increasing the sensitivity of a gravitational wave detector. A standard bar detector for gravitational radiation may be treated as a harmonic oscillator. The effect of the gravitational radiation is so weak that the expected displacement of the bar is of the order of 10^{-19} cm. This is the same order of magnitude as the quantum mechanical uncertainty of the bar's position in its ground state. Thus the signal from the gravitational wave detector may be obscured by the zero-point fluctuations of the detector. This is a striking example of the influence of quantum fluctuations on a macroscopic system. In principle, a way of beating this problem is clear. Instead of the ground state of the oscillator with its quantum noise randomly distributed in phase one prepares the oscillator in a squeezed state. One then measures the displacement due to the gravitational radiation in the quadrature with reduced fluctuations. In this way it should be possible to detect displacements less than the quantum mechanical uncertainty in the bar's position. Of course, this leaves a lot of technical questions unanswered. How does one prepare the bar in a squeezed state? How does one make a measurement on the bar's quadrature phase? These problems and suggested solutions are discussed elsewhere^{8,9} in treatments of quantum non-demolition measurements.

The statistical properties of light fields such as coherent or thermal light may be calculated by techniques similar to classical

probability theory using an expansion of the density operator in terms of coherent states, the Glauber-Sudarshan P representation^{10,11}. Coherent light has poissonian photon counting statistics. Squeezed states of light on the other hand may have sub-poissonian photon counting statistics and have no nonsingular representation in terms of the Glauber-Sudarshan P distribution. The statistical properties of such fields cannot be calculated by techniques analogous to classical probability theory. Squeezed states are, therefore, an example of a nonclassical light field. To be precise we shall define a nonclassical light field as one that has no positive nonsingular Glauber-Sudarshan P function.

Another example of a nonclassical light field is a number state. This certainly has no nonsingular Glauber-Sudarshan P function and clearly has sub-poissonian photon statistics. Such nonclassical light fields with sub-poissonian photon statistics which exhibit photon antibunching have been observed experimentally^{12,13,50}. A number state, however, has its quantum fluctuations randomly distributed in phase and hence does not exhibit squeezing. While a squeezed state may exhibit sub-poissonian photon statistics and hence photon antibunching it is not a necessity. Sub-poissonian statistics result if the quadrature phase with reduced fluctuations carries the coherent excitation. Using photon counting techniques direct measurements of the intensity fluctuations of a light field are possible. To determine the fluctuations in the quadrature phases a phase sensitive detection scheme is necessary. This can be achieved by homodyning or heterodyning the signal with a local oscillator followed by photon counting measurements. To generate a squeezed state a phase dependent nonlinear optical process is necessary.

Phase dependent correlation functions

Detection of a light signal with a photon counter yields a measurement of the light intensity $I(t)$ or photon number $n(t)$. Using electronic correlators one may then compute the intensity or photon number correlations of the light field. For example, one may measure the normalized second-order correlation function

$$g^{(2)}(0) = \frac{\langle :I^2: \rangle}{\langle I \rangle^2} \quad (1)$$

where $::$ denotes normal ordering of the quantum mechanical operators. For sufficiently short counting times the variance $V(n)$ of the photon number distribution is related to $g^{(2)}(0)$ by

$$\frac{V(n) - \langle n \rangle}{\langle n \rangle} = g^{(2)}(0) - 1 \quad (2)$$

A coherent light field with poissonian statistics has $g^{(2)}(0) = 1$. Thermal light which has increased intensity fluctuations has $g^{(2)}(0) = 2$. Since $g^{(2)}(0)$ represents the probability of two photons arriving simultaneously this is referred to as photon

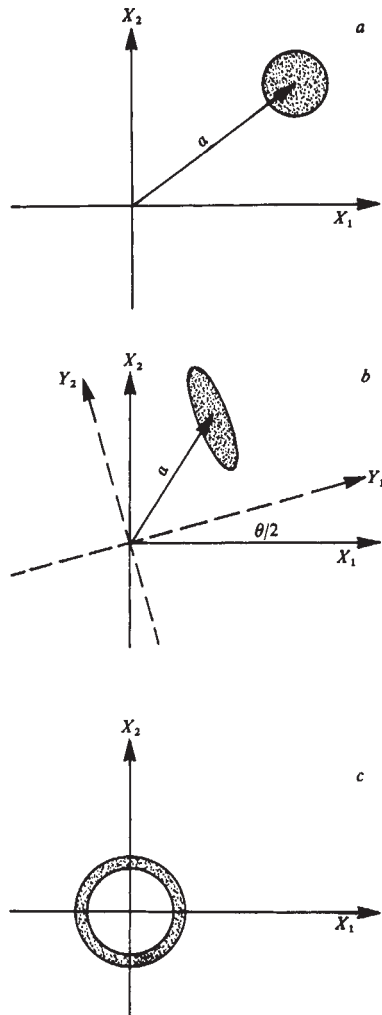


Fig. 1 Phase space plot showing the uncertainty in: *a*, a coherent state $|\alpha\rangle$; *b*, a squeezed state $|\alpha, r e^{i\theta}\rangle$ ($r > 0$); *c*, a number state $|n\rangle$.

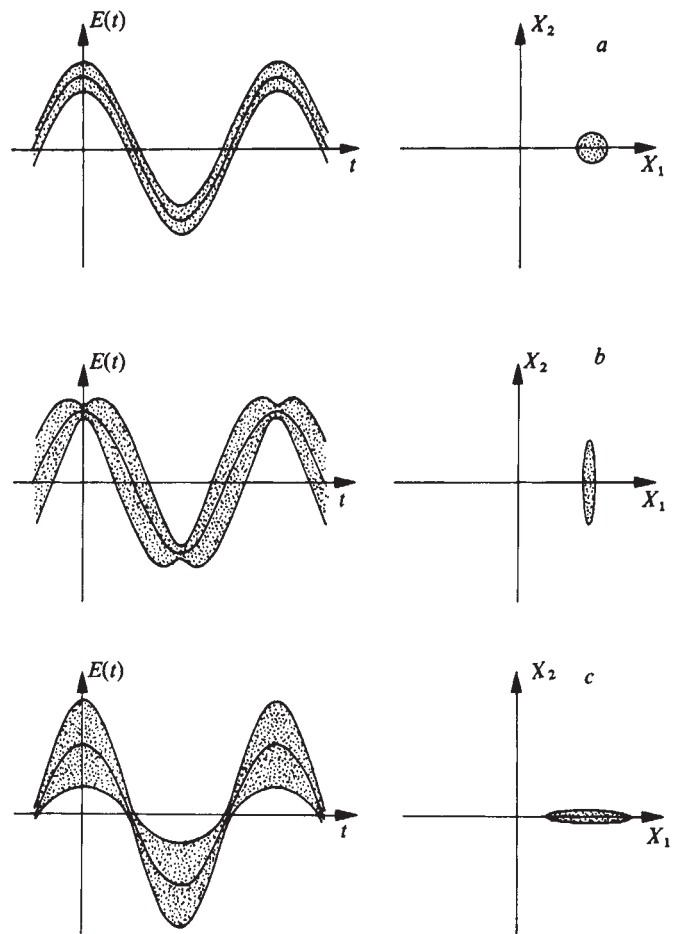


Fig. 2 Plot of electric field against time showing the uncertainty for: *a*, a coherent state $|\alpha\rangle$ (α real); *b*, a squeezed state $|\alpha, r\rangle$ with reduced amplitude fluctuations (α real, $r > 0$); *c*, a squeezed state $|\alpha, r\rangle$ with reduced phase fluctuations (α real, $r < 0$). Reproduced with permission from Caves²¹.

bunching. A light field with sub-poissonian statistics will have $g^{(2)}(0) < 1$, an effect known as photon antibunching. Photon antibunching is a quantum mechanical effect which may not be derived from a classical description of the field. Such fields do not have a positive nonsingular representation in terms of the Glauber–Sudarshan P distribution which expresses the density operator for a single mode field as^{10,11}

$$\rho = \int P(\alpha) |\alpha\rangle \langle \alpha| d^2\alpha \quad (3)$$

where $|\alpha\rangle$ is a coherent state. This representation has found considerable application in optics because the taking of quantum mechanical averages resemble classical averaging procedures provided $P(\alpha)$ exists as a positive nonsingular function. For fields which exhibit photon antibunching, however, the $P(\alpha)$ are highly singular functions. In this sense we say that such fields are nonclassical. The quantum theory of light received further verification when photon antibunching was observed experimentally in resonance fluorescence from a two level atom^{12,13} in agreement with theoretical predictions^{14–16} (for reviews see refs 17–19).

Our discussion of the properties of phase dependent correlation functions is illustrated with reference to a single mode field. We may write the electric field as

$$E(t) = \lambda (a e^{-i\omega t} + a^\dagger e^{i\omega t}) \quad (4)$$

where λ is a constant including the spatial wave functions. In the quantum theory of radiation the amplitudes a and $a^\dagger$ are

quantum mechanical operators which obey boson commutation relations. We may write

$$a = X_1 + iX_2 \quad (5)$$

where X_1 and X_2 are hermitian operators obeying the commutation relation

$$[X_1, X_2] = \frac{i}{2} \quad (6)$$

In terms of X_1 and X_2 one may write $E(t)$ as

$$E(t) = \frac{\lambda}{2} (X_1 \cos \omega t + X_2 \sin \omega t) \quad (7)$$

Thus X_1 and X_2 may be identified as the amplitudes of the two quadrature phases of the field.

From the commutation relation (6) we deduce the following relation for the uncertainties $\Delta X_i = \{V(X_i)\}^{1/2}$ in X_1 and X_2

$$\Delta X_1 \Delta X_2 \geq \frac{1}{4} \quad (8)$$

A family of minimum uncertainty states is defined by taking the equal sign. One such class of minimum uncertainty states is the coherent states which have $V(X_1) = V(X_2) = \frac{1}{4}$. A broader class of minimum uncertainty states may have unequal variances in each quadrature. These are the so called squeezed states. The condition for squeezing is

$$V(X_i) < \frac{1}{4} \quad i = 1 \text{ or } 2 \quad (9)$$

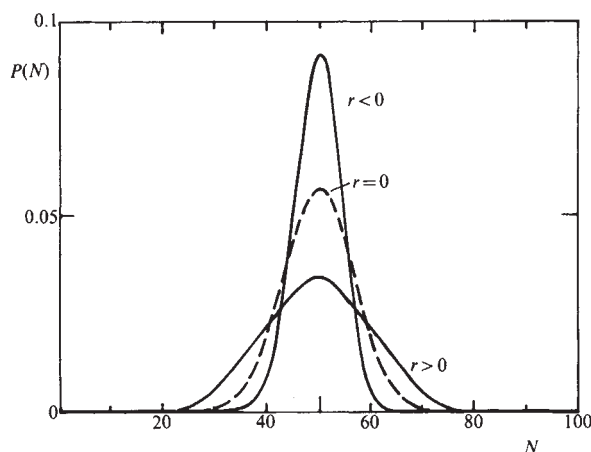


Fig. 3 Photon number distribution for a squeezed state $|\alpha, r\rangle$ ($\alpha = 7$, $r = \pm 0.5$) compared with a coherent state ($r = 0$).

It is sometimes convenient especially for multimode fields to write the condition in terms of the normally ordered variance

$$:V(X_i): < 0 \quad i = 1 \text{ or } 2 \quad (10)$$

Figure 1 shows a phase space plot of the uncertainties in X_1 and X_2 for a coherent state, a squeezed state and a number state is shown. These error ellipses may be rigorously derived as the contours of the Q function⁴.

The time dependence of $E(t)$ including the uncertainty $\Delta E(t)$ is shown in Fig. 2 for a a coherent state, b a squeezed state with reduced amplitude fluctuations, c a squeezed state with reduced phase fluctuations. The corresponding error box for these states at $t = 0$ is also shown.

For a single mode field the variance in one quadrature may be calculated using the Glauber-Sudarshan P representation

$$V(X_1) = \frac{1}{4} \left\{ 1 + \int P(\alpha) [(\alpha + \alpha^*) - (\langle \alpha \rangle + \langle \alpha^* \rangle)]^2 d^2 \alpha \right\} \quad (11)$$

The condition for squeezing $V(X_1) < \frac{1}{4}$ requires that $P(\alpha)$ be a nonpositive definite function. In this sense squeezing like photon antibunching is a nonclassical property of the electromagnetic field. Note that to derive equation (11) the commutation relation $[a, a^\dagger] = 1$ has been used. If a classical field is assumed from the outset arbitrary squeezing may be obtained in either quadrature. Thus squeezing has a non trivial significance only in the case of quantized fields. A distinction between classical and quantum fields may be obtained from the normally ordered correlation function $g^{(2)}(0)$ which is always ≥ 1 for classical fields (see ref. 20).

Properties of squeezed states

We shall now briefly describe the mathematical properties of squeezed states. A coherent state $|\alpha\rangle$ may be generated by the action of the displacement operator $D(\alpha)$ on the vacuum

$$|\alpha\rangle = D(\alpha)|0\rangle \quad (12)$$

where

$$D(\alpha) = e^{-\frac{1}{2}|\alpha|^2} e^{a\alpha^\dagger} e^{-a^* \alpha}$$

A squeezed state $|\alpha, \zeta\rangle$ may be generated by first acting with the squeeze operator $S(\zeta)$ on the vacuum followed by the displacement operator $D(\alpha)$ (ref. 21)

$$|\alpha, \zeta\rangle = D(\alpha)S(\zeta)|0\rangle \quad (13)$$

where

$$S(\zeta) = \exp\left(\frac{1}{2}\zeta^* a^2 - \frac{1}{2}\zeta a^{\dagger 2}\right)$$

and

$$\zeta = r e^{i\theta}$$

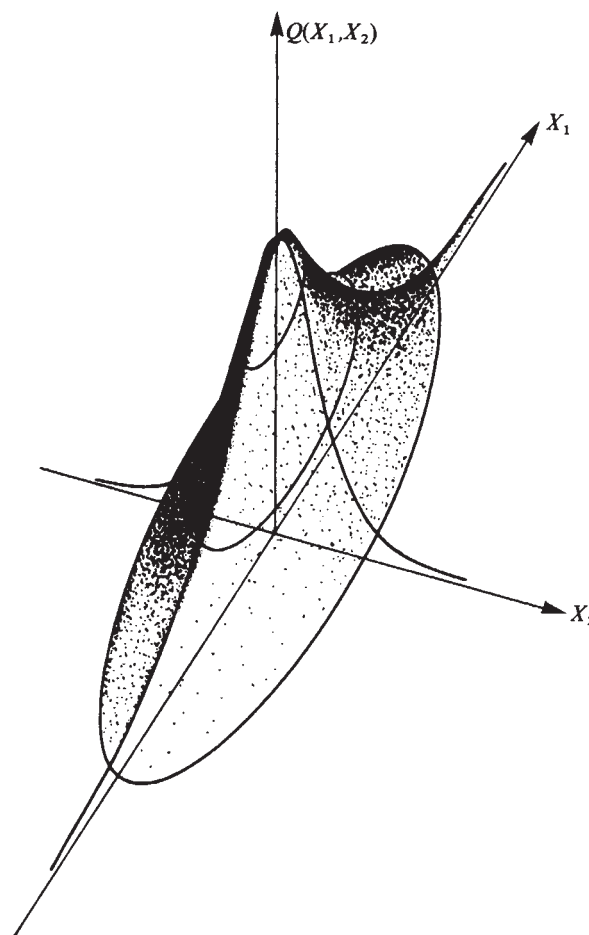


Fig. 4 Q function for a squeezed state. Reproduced with permission from Yuen⁴.

An alternative but equivalent characterization of squeezed states has been given by Yuen⁴. We note that whereas a coherent state is generated by linear terms in a and $a^\dagger$ in the exponent a squeezed state requires quadratic terms.

The variances in squeezed state $|\alpha, \zeta\rangle$ are given by

$$\begin{aligned} V(Y_1) &= \frac{1}{4} e^{-2r} \\ V(Y_2) &= \frac{1}{4} e^{2r} \end{aligned} \quad (14)$$

where $Y_1 + iY_2 = (X_1 + iX_2) e^{-i\theta/2}$ is a rotated complex amplitude so that $2\Delta Y_1$ and $2\Delta Y_2$ represent the length of the minor and major axes of the error ellipse. The mean photon number in the squeezed state $|\alpha, \zeta\rangle$ is

$$\langle n \rangle = |\alpha|^2 + \sinh^2 r \quad (15)$$

Clearly, the variances $V(X_i)$ are independent of the field amplitude α . Thus squeezing is a quantum mechanical effect which may occur in fields with high intensity. In this sense one may say it is a macroscopic quantum effect. This is a significant difference from photon antibunching which is only appreciable for fields with low intensity. There is no general relation between photon antibunching and squeezing, however, we shall consider the limit where the coherent amplitude greatly exceeds the squeezing ($|\alpha|^2 \gg \sinh^2 r$). In this limit we find

$$:V(X_1): = \frac{\alpha^2}{4} (g^{(2)}(0) - 1) = \frac{1}{4} (e^{-2r} - 1) \quad (16)$$

where we have chosen α real so that the amplitude is carried by X_1 .

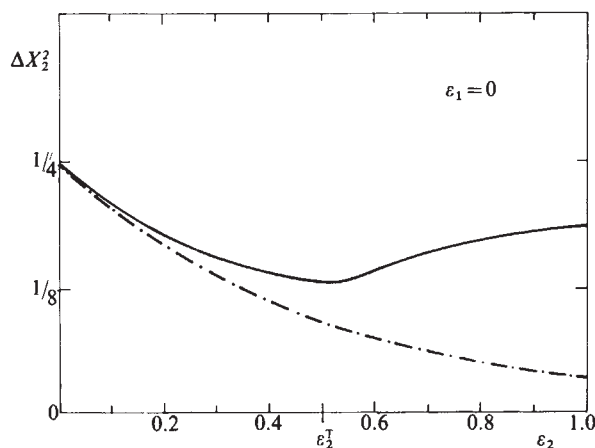


Fig. 5 Squeezing in the parametric oscillator (—) compared with an ideal parametric amplifier (---) as a function of the pump driving field.

For $r > 0$ we have a reduction in amplitude fluctuations ($V(X_1) < 0$) and photon antibunching ($g^{(2)}(0) < 1$) whereas for $r < 0$ we have an increase in amplitude fluctuations and photon bunching. Hence in this limit a squeezed state may show either photon bunching or antibunching depending on whether the amplitude fluctuations are increased or reduced. The photon number distribution⁴ in a squeezed state $|\alpha, r\rangle$ is plotted in Fig. 3 for $\alpha = 7$, $r = \pm 0.5$. We can see that the photon statistics are sub- or super-poissonian depending on whether $r > 0$ or $r < 0$.

No such simple relation between antibunching and squeezing exists for all values of α . For example in the opposite limit of $\alpha \ll 1$, that is, a squeezed vacuum $|0, r\rangle$,

$$g^{(2)}(0) = 1 + \frac{\cosh 2r}{\sinh^2 r} \quad (17)$$

Thus the photons in a squeezed vacuum are always bunched irrespective of the sign of the squeeze parameter.

An example of a quantum state which exhibits photon antibunching but no squeezing is a number state $|n\rangle$ for which

$$V(X_1) = V(X_2) = \left(\frac{1}{2}\right)(2n+1) \quad (18)$$

The complete absence of phase information in a number state is clear from the phase space annulus shown in Fig. 1c.

As the Glauber-Sudarshan P representation does not exist for a squeezed state we must consider an alternative representation such as the Wigner function, the Q function, or the generalized P function^{22,23}. The Q function for a squeezed state is derived in ref. 4. The distribution $Q(X_1, X_2)$ plotted in Fig. 4 as a function of the amplitudes of the two quadratures clearly shows the unequal variances in X_1 and X_2 .

Production of squeezed states

There has been no experimental manifestation of squeezed states of light. The requirement to produce a squeezed state may be simply expressed as follows. For a single mode field mix a part of the field with its phase conjugate to produce a new mode b such that

$$b = \mu a + \nu a^\dagger \quad (19)$$

where $\mu^2 - \nu^2 = 1$. For mode a in a coherent state the mode b will be in a squeezed state⁴. Thus a scheme involving a phase

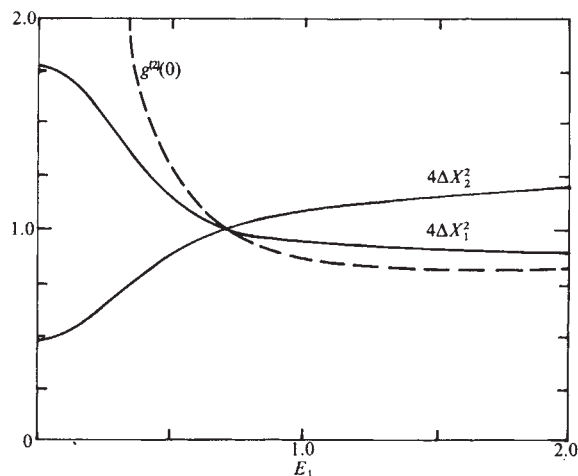


Fig. 6 Photon statistics $g^{(2)}(0)$ and variances ΔX_i^2 for the parametric oscillator as a function of the idler driving field. The pump driving field is held fixed at the threshold value.

conjugate mirror appears as one of the favourite candidates for a state squeezer²⁴. The above prescription seems very simple, however, the phase conjugate mirror involves a nonlinear interaction, in this case a four-wave mixing interaction. Squeezed states may also be generated by a three-wave mixing interaction as for example in the parametric amplifier^{25-27,51,52}. The prototype for these interactions is described by the Hamiltonian,

$$H = \hbar[\chi^{(n)*}(\epsilon)a^2 + \chi^{(n)}(\epsilon)a^{\dagger 2}] \quad (20)$$

where

$$\chi^{(2)}(\epsilon) = \epsilon \chi^{(2)} \quad (\text{degenerate parametric amplifier})$$

$$\chi^{(3)}(\epsilon) = \epsilon^2 \chi^{(3)} \quad (\text{four-wave mixing})$$

$\chi^{(n)}$ is the nonlinear susceptibility of the optical medium and ϵ is the amplitude of the pump field which has been treated classically. This approximate form of the Hamiltonian generates squeezed states with a squeeze parameter given by $r = |2\chi^{(n)}(\epsilon)t|$. Several objections to this ideal system can be raised. Fluctuations resulting from the quantization of the pump field and the nonlinear medium have been neglected as have vacuum fluctuations associated with any loss process. The vacuum fluctuations will tend to equalize the variances in the two quadratures and hence destroy the squeezing²⁸⁻²⁹. Thus the characteristic damping time of any loss mechanism should be long compared with the interaction time. Phase and amplitude fluctuations in the laser used for the pump may also degrade the squeezing³⁰. Phase fluctuations may be compensated for by using part of the pump as the local oscillator in a homodyne detection scheme.

The magnitude of the squeezing is limited by the small values of the nonlinear susceptibility and the interaction time. To increase the interaction time the nonlinear crystal may be placed inside an optical cavity. There is a parametric oscillator configuration where the cavity modes are driven externally by classical fields. An analysis of the cavity must include the cavity losses which tend to destroy the squeezing. Thus there will be a competition between the squeezing produced by the nonlinear interaction and the degradation of the squeezing by the damping. This results in a limiting value to the squeezing attainable in the steady state. An analysis of the degenerate parametric oscillator including the quantization of the pump field has been carried out³¹⁻³³. When only the pump mode is driven by an external field there exists a threshold driving field below which the semiclassical value of the mean field is zero. The squeezing in the idler mode as a function of the pump field amplitude is

shown in Fig. 5. As the pump amplitude is increased from zero squeezing appears in the idler mode. However, the squeezing approaches a maximum value corresponding to $V(X_2) = 1/8$ close to the threshold value of the pump field, then decreases as the pump power is increased above threshold.

The case where the driving field for the pump mode is held fixed at the threshold value and the driving field for the idler is increased from zero is shown in Fig. 6. For low values of the idler driving field the squeeze parameter is initially positive and amplitude fluctuations are increased, hence we have photon bunching ($g^{(2)}(0) > 1$). As the idler driving field is increased the squeeze parameter goes to zero and then becomes negative. Thus we have reduced amplitude fluctuations and hence photon antibunching ($g^{(2)}(0) < 1$). This is consistent with the general properties of squeezed states with $|\alpha|^2 \gg \sinh^2 r$ discussed above. This system provides a feasible scheme for detecting squeezed states by making photon correlation measurements directly on the squeezed field. Facility to change the sign of the squeeze parameter and observe the accompanying change of the photon statistics from bunching to antibunching would indicate the presence of a squeezed state.

Other nonlinear intracavity devices have been shown^{34,35} to give a maximum squeezing factor not greatly exceeding 2. The coupling of the cavity modes to the vacuum fluctuations of the extracavity modes apparently acts as a counter to the squeezing produced by the nonlinear interaction in a steady-state configuration.

One possibility of avoiding the limitation to squeezing imposed by the vacuum fluctuations entering the cavity is to make one mirror perfectly reflecting. It has been claimed that since the vacuum fluctuations may no longer enter from the second port arbitrary squeezing is in principle attainable (B. Yurke, personal communication).

Another way to avoid the problem of vacuum fluctuations is to revert to the parametric amplifier configuration where the cavity losses no longer have a role. The parametric amplifier is a travelling wave phase matched interaction and the Hamilton equation (20) which only includes a single mode is not appropriate. A multimode analysis³⁶ of a travelling wave parametric amplifier indicates that a reduction in squeezing over the single mode case may occur for the non degenerate amplifier. This reduction in squeezing is caused by the contribution from non-resonant modes whose axes of squeezing become misaligned with respect to the resonant mode.

Another possible system for producing squeezed states is a two-photon laser due to the quadratic nature of the field interaction. A laser, however, is an active system in which the atoms are pumped to the excited state and may consequently decay by spontaneous emission. Calculations using a two-level model for the atomic medium reveal that any potential squeezing is destroyed by the fluctuations resulting from spontaneous emission^{37,38}.

It is clear, therefore, that a phase sensitive nonlinear interaction in a passive medium is required to produce squeezed states. Predictions of squeezing in a variety of nonlinear optical processes have now been made, for example the free electron laser³⁹, second harmonic generation^{40,41}, the single atom-single field mode interaction⁴², and multiphoton absorption⁴³. The prediction of squeezing in four wave mixing²⁴ has attracted the interest of experimentalists⁴². An analysis of the effect of atomic fluctuations in four-wave mixing based on a two-level atomic medium reveals that for the atoms driven near saturation or close to resonance the spontaneous emission will destroy the squeezing⁴⁵. For significant squeezing the driving fields should be of low intensity and sufficiently far from resonance so as not to saturate the atoms.

Another system with somewhat different characteristics is squeezing in resonance fluorescence from a two-level atom⁴⁶. Resonance fluorescence differs from many of the systems discussed above as it involves many modes of the radiation field. Resonance fluorescence deserves attention as it is the only system in which photon antibunching has been observed^{12,13}.

We consider a two-level atom driven by a coherent driving field. The product of the amplitude of the driving field and the dipole moment of the atom is characterized by the Rabi frequency. We denote the Rabi frequency normalized by the natural linewidth of the atom by Ω . The driving field may have a detuning with respect to the atomic transition. We shall use δ to characterize the detuning normalized by the natural linewidth.

The condition for squeezing in a field may best be expressed in terms of the normally ordered variances which do not include the contribution from the vacuum fluctuations. For squeezing in either quadrature ($E_1 = (E^{(+)} + E^{(-)})/2$, $E_2 = (E^{(+)} - E^{(-)})/2i$) of the field we require

$$:V(E_i): < 0 \quad i = 1 \text{ or } 2 \quad (21)$$

We calculate the squeezing in the components of the fluorescent field in the direction along and perpendicular to the mean field. The variance in the component E'_1 in the direction of the mean field is

$$:V(E'_1): = \frac{-\lambda}{4} \frac{(1 + \delta^2 - \Omega^2)}{1 + \delta^2 + \Omega^2} \Omega^2 \quad (22)$$

where λ is a constant.

Thus we find squeezing in this component provided $\Omega^2 < 1 + \delta^2$. No squeezing occurs in the component orthogonal to the mean field. The reduced amplitude fluctuations occurring for $\Omega^2 < 1 + \delta^2$ is consistent with the observed fact that the fluorescent light is antibunched. We note that the fluorescent light is also antibunched in the strong field limit $\Omega^2 > 1 + \delta^2$ where there is no squeezing. In this limit the characteristics of the fluorescent light resemble a number state.

Detection of squeezed states

Proposals to measure the variances in the quadrature phases of a light field suggest homodyning or heterodyning the signal with a local oscillator which gives the necessary phase dependence followed by a photon counting measurement. Such measurements are feasible with existing technology. (For further details of such a measurement scheme see refs 6, 7, 20.)

The signal field is homodyned with a local oscillator which is assumed to be in a coherent state. The complex amplitude of the local oscillator may be written as $\varepsilon = |\varepsilon| e^{i\theta}$ where θ is the phase of the local oscillator with respect to the signal field. In the limit where the amplitude of the local oscillator greatly exceeds the amplitude of the signal field the photon statistics of the combined field are directly related to the normally ordered variance of the signal field. Assuming a perfect detector efficiency it may be shown that^{6,7,20}

$$\begin{aligned} V(n) - \langle n \rangle &= 4|\varepsilon|^2 :V(E_1): \quad \text{if } \theta = 0 \\ &= 4|\varepsilon|^2 :V(E_2): \quad \text{if } \theta = \pi/2 \end{aligned} \quad (23)$$

Thus by changing the phase of the local oscillator a measurement of the photon statistics yields the normally ordered variance in E_1 ($\theta = 0$) and the normally ordered variance in E_2 ($\theta = \pi/2$). A change of photon statistics from sub- to super-poissonian as θ is varied will indicate the presence of squeezing.

Such measurements impose a stringent requirement on the relative phase stability between the local oscillator and the signal. Yuen and Chan⁴⁷ have recently suggested that photon number fluctuations in the local oscillator may be eliminated using a balanced detector scheme developed in the microwave region⁴⁸.

Another way to detect a squeezed state is by a direct photon correlation measurement if one has the facility to vary the sign of the squeeze parameter. The presence of a squeezed state is

indicated by a change of photon statistics from bunching to antibunching as the squeeze parameter is varied. An example of such a system is the parametric oscillator with two driving fields discussed earlier. This method obviates the need for homodyning the signal with a local oscillator.

Applications of squeezed states

Squeezed states have several potential applications, one, for example, is in optical communication systems. In a proposed scenario information would be transmitted in the quadrature of the field with reduced quantum fluctuations. An enhanced signal-to-noise ratio could then be obtained in the quantum noise limited regime over information sent using coherent light beams. The application of squeezed states in optical communications systems is discussed in refs 6 and 7.

Similar considerations hold in the amplification of signals. Noise is necessarily added in the amplification process, however, if a suitable phase sensitive amplifier is used the noise may be added preferentially to the quadrature not carrying information. This leaves the amplification of the quadrature carrying the information essentially noise free.

Interferometric techniques to detect very weak forces such as gravitational radiation experience limitations on sensitivity due to quantum noise arising from photon counting and radiation pressure fluctuations. These sources of noise may be interpreted as arising from the beating of the input laser with the vacuum fluctuations entering the unused port of the interferometer. It turns out that these two different noise sources arise from fluctuations in the two different quadrature phases of the vacuum entering the unused input port. It has been suggested by Caves²¹ that injecting a squeezed state into the unused input port will reduce one or other of the two sources of noise depending on which quadrature is squeezed.

Another intriguing application of squeezed states is in an optical waveguide tap. Shapiro has shown that a high signal-to-noise ratio may be obtained using a squeezed state in an optical waveguide to tap a signal carrying waveguide⁴⁹. This may be achieved with very low energy loss from the signal thus offering the possibility of permitting optical data bus technology to reach multikilometre path lengths with many user sites but no repeaters.

Conclusions

The field of quantum optics has been an active field of research since the early 1960s. However, much which has been discussed

under this heading could more correctly be described as non-linear optics as no quantization of the electromagnetic field is necessary. Very few features which are explicitly a result of quantization of the field have been observed—photon antibunching being one exception. Squeezed states represent a class of quantum states for which no classical analogue exists, hence their detection would be of fundamental interest.

The achievements of quantum optics have been based on the measurement of photon correlation functions of the electromagnetic field. We now seem to be on the verge of an era where a new class of measurements on the phase dependent correlation functions of the electromagnetic field will be possible. This will enable information on the electromagnetic field to be obtained which was not accessible from photon correlation measurements. Such measurements based on homodyning or heterodyning the field with a local oscillator appear feasible with current technology. The presence of a squeezed state will be indicated by the observation of sub-poissonian photon statistics in such a phase sensitive detection process.

Present efforts are directed towards methods of generating a squeezed state. While a proof in principle of the existence of squeezed states seems possible in, for example, resonance fluorescence from a two-level atom or an intracavity nonlinear optical interaction the magnitude of the squeezing obtained in such systems is small. To obtain appreciable squeezing one must look to either a single pass device with a high nonlinearity and low losses or possibly to a cavity with a single input/output port which prevents the vacuum fluctuations entering as in the two port cavity.

Should a device be found to give a light field with significant squeezing the potential applications are attractive. These applications lie on the frontier of technology in quantum noise limited situations. For example, a squeezed light field could be used in an optical communication system where the information is carried by the quadrature with reduced quantum fluctuations. This would enable a better signal-to-noise ratio to be attained than using conventional laser sources which are limited by the quantum noise of a coherent state. The general concept of squeezed states with their phase dependence of quantum noise has important implications in quantum amplifier theory and ultrasensitive electronics such as required for the detection of gravitational radiation. While no experimental observation of squeezed states has yet been reported this is a goal well worth achieving both from a fundamental point of view and in consideration of the applications that will follow.

I thank H. J. Carmichael and G. J. Milburn for valuable suggestions. This work was supported in part by the United States Army through its European Research Office.

- Robinson, D. R. *Commun. Math. Phys.* **1**, 159 (1965).
- Stoler, D. *Phys. Rev. D* **1**, 3217 (1970); **D4**, 1925 (1971).
- Lu, E. Y. C. *Lett. Nuovo Cimento* **2**, 1241 (1971); **4**, 585 (1972).
- Yuen, H. P. *Phys. Rev. A* **13**, 2226 (1976).
- Hollenhorst, H. N. *Phys. Rev. D* **19**, 1669 (1979).
- Yuen, H. P. & Shapiro, J. H. *IEEE Trans. Inform. Theory* **IT24**, 657 (1978); **IT26**, 78 (1980).
- Shapiro, J. H., Yuen, H. P. & Machado Mata, J. A. *IEEE Trans. Inform. Theory* **IT25**, 179 (1979).
- Braginsky, V. B., Vorontsov, Yu. I. & Thorne, K. S. *Science* **209**, 547 (1980).
- Caves, C. M., Thorne, K. S., Drever, R. W. P., Sandberg, V. D. & Zimmerman, M. Rev. *Mod. Phys.* **52**, 341 (1980).
- Glauber, R. J. *Phys. Rev.* **131**, 2766 (1963).
- Sudarshan, E. C. G. *Phys. Rev. Lett.* **10**, 277 (1963).
- Kimble, H. J., Dagenais, M. & Mandel, L. *Phys. Rev.* **18A**, 201 (1978).
- Leuchs, G., Rateike, M. & Walther, H. (in preparation).
- Carmichael, H. J. & Walls, D. F. *J. Phys.* **9B**, 1199 (1976).
- Kimble, H. J. & Mandel, L. *Phys. Rev. A* **13**, 2123 (1976).
- Cohen Tannoudji, C. in *Frontiers in Laser Spectroscopy* (eds Balian, R., Haroche, S. & Liberman, S.) (North Holland, Amsterdam, 1977).
- Walls, D. F. *Nature* **280**, 451 (1979).
- Loudon, R. *Rep. Prog. Phys.* **43**, 58 (1958).
- Cresser, J. D., Häger, J., Leuchs, G., Rateike, M. & Walther, H. in *Dissipative Systems in Quantum Optics* (ed. Bonifacio, R.) **21** (Springer, Berlin, 1982).
- Mandel, L. *Phys. Rev. Lett.* **49**, 136 (1982).
- Caves, C. M. *Phys. Rev.* **23D**, 1693 (1981).
- Drummond, P. D. & Gardiner, C. W. *J. Phys.* **13A**, 211 (1980).
- Drummond, P. D., Gardiner, C. W. & Walls, D. F. *Phys. Rev.* **24A**, 914 (1981).
- Yuen, H. P. & Shapiro, J. H. *Opt. Lett.* **4**, 334 (1979).
- Takahashi, H. *Adv. Commun. Syst.* **1**, 227 (1965).
- Raiford, M. T. *Phys. Rev. A* **2**, 1541 (1970); **A9**, 2060 (1974).
- Mista, L., Perinova, V., Perina, J. & Braunerova, Z. *Acta Phys. Pol.* **A51**, 739 (1977).
- Hillery, M. & Scully, M. O. in *Quantum Optics, Experimental Gravitation and Measurement Theory* (eds Meystre, P. & Scully, M. O.) (Plenum, New York, in the press).
- Milburn, G. & Walls, D. F. *Am. J. Phys.* (in the press).
- Wodkiewicz K. & Zubairy, M. S. *Phys. Rev.* **27A**, 2003 (1983).
- Milburn, G. J. & Walls, D. F. *Opt. Commun.* **39**, 401 (1981).
- Lugiato, L. A. & Strini, G. *Opt. Commun.* **41**, 67 (1982).
- Milburn, G. J. & Walls, D. F. *Phys. Rev.* **27A**, 392 (1983).
- Walls, D. F. & Milburn, G. J. in *Quantum Optics, Experimental Gravitation and Measurement Theory* (eds Meystre, P. & Scully, M. O.) (Plenum, New York, in the press).
- Lugiato, L. A. & Strini, G. *Opt. Commun.* **41**, 447 (1982).
- Tombesi, P., Lane, A., Carmichael, H. J. & Walls, D. F. *Opt. Commun.* (in the press).
- Lugiato, L. A. & Strini, G. *Opt. Commun.* **41**, 374 (1982).
- Reid, M. D. & Walls, D. F. *Phys. Rev.* **28A**, 332 (1983).
- Becker, W., Scully, M. O. & Zubairy, M. S. *Phys. Rev. Lett.* **48**, 475 (1982).
- Mandel, L. *Opt. Commun.* **42**, 437 (1982).
- Lugiato, L. A., Strini, G. & de Martini, F. *Opt. Lett.* **8**, 256 (1983).
- Meystre, P. & Zubairy, M. S. *Phys. Lett.* **89A**, 390 (1982).
- Zubairy, M. S., Razmi, M. S. K., Iqbal, S. & Idrees, M. (in preparation).
- Kumar, P., Bondurant, R. S., Shapiro, J. H. & Salour, M. M. *Proc. 5th Rochester Conf. on Coherence and Quantum Optics* (eds Mandel, L. & Wolf, E.) (Plenum, New York, in the press).
- Reid, M. D. & Walls, D. F. *Opt. Commun.* (in the press).
- Walls, D. F. & Zoller, P. *Phys. Rev. Lett.* **47**, 709 (1981).
- Yuen, H. P. & Chan, V. W. S. *Opt. Lett.* **8**, 177 (1983).
- Dicke, R. H. *Rev. Scient. Instrum.* **17**, 268 (1946).
- Shapiro, J. H. *Opt. Lett.* **5**, 351 (1980).
- Short, R. & Mandel, L. *Phys. Rev. Lett.* **51**, 384 (1983).
- Mollow, B. R. & Glauber, R. J. *Phys. Rev.* **160**, 1097 (1967).
- Mollow, B. R. *Phys. Rev.* **162**, 1256 (1967).

Fault related folding near the Wind River thrust, Wyoming, USA

Geoffrey King & Jonathan Brewer

Bullard Laboratories, Department of Earth Science, University of Cambridge, Madingley Road, Cambridge CB3 0EZ, UK

Reflection profiles across the Wind River uplift in Wyoming, USA, show that folding of basement rocks occurs in the footwall of the main fault at depths between 6 and 10 km. It is shown, using elastic modelling, that these features are a consequence of uneven fault slippage. Slip on the fault must be greater at depths between 3 and 18 km, but reduced outside this range. This is the depth range of catastrophic rupture associated with earthquakes in the continental crust.

COCORP (Consortium for Continental Reflection Profiling) data recorded across the Wind River mountains, one of the Laramide basement uplifts of the Cordilleran Foreland of the western USA, have successfully located a major thrust fault extending to middle and lower crustal depths (Fig. 1)^{1,2}. The Precambrian crystalline core of the mountains has overridden sedimentary rocks of the Green River basin by a horizontal distance of at least 17.5 km along a surface dipping at between 30° and 40° and striking at approximately 140°. The minimum vertical offset of the basement corresponding to the geometry identified by the profiles is 9 km. In the hanging wall of the thrust, at a distance of about 30 km from the thrust front, Phanerozoic sedimentary rocks dip at about 10° off the backside of the mountains into the Wind River Basin.

It has long been recognized that these mountains result from a series of basement uplifts of late Cretaceous–early Tertiary age³. The style of deformation is in marked contrast to the thin-skinned deformation that has occurred elsewhere in the Cordilleran Foreland, and much controversy has centred on whether the basement uplifts were driven by vertical or horizontal crustal movements. The COCORP data resolved this question, for the Wind River mountains at least, in favour of horizontal motion being converted to vertical motion as a result of movement on deep-seated, relatively shallow angle thrust faults. In addition, they have revealed fascinating features of the deformation of the basement and lower sedimentary rocks in the footwall of the thrust. Folding of the deepest sedimentary layers and possible reverse faulting can be seen and has been interpreted as the initial expression of the shortening process, preceding the development of the thrust plane^{2,4}. Here we show that the folding is more likely to be contemporary with, and as a consequence of, the thrust fault motion and is not due to layer parallel shortening. Our understanding of the mechanics of fault related folding, and of modelling methods, comes from studies of the deformation associated with recent large earthquakes. The significance of these studies for understanding the Wind River thrust are now discussed.

Work on the Tabas (Iran) earthquake⁵ and the El Asnam (Algeria) earthquake^{6,7} demonstrated that surface anticlines associated with active thrust faults increased in amplitude synchronously with seismic fault motion at depth. In Algeria the nature of the folding process can be well constrained by a wealth of geomorphological, geodetic, geological and seismic information. Changes of drainage and geodetic data⁸ demonstrated an increase of amplitude of the Sera el Maarouf anticline of about 5 m at the time of the earthquake. River terraces and old survey maps⁹ indicated that similar motions had occurred in historical times and carbon dates⁹ and archaeological artefacts in tilted gravels⁶ have been used to outline a chronology of fold development. The foregoing information suggests that the Sera el Maarouf anticline results from the superimposition of the deformation

of about 30 earthquakes similar to the recent one, occurring over a period of about 100,000 yr. Seismic records of that earthquake together with aftershock studies provide information about the amplitude and orientation of motion on the buried fault and suggest that the development of the fold could be explained by adapting current ideas of catastrophic rupture propagation.

The environment of the fault of a shallow earthquake in a continental environment is shown in Fig. 2a. An upper zone nominally 5 km thick is low in stress before the earthquake. This may be attributed to creep promoted by the chemical effects of crack and pore water in an environment of low confining pressure. Note that if stress is relieved by stress corrosion cracking along joints, the system is only low in stress in an average sense. Blocks between joints may have substantial internal stresses. Beneath the nominal 5 km depth, the fissures close sufficiently to suppress creep mechanisms enhanced by the presence of water in fissures or pores and in a depth interval nominally between 5 and 20 km, stress can rise to failure level as a result of tectonic loading. Below 20 km thermally activated crystalline creep processes¹⁰ relieve stress faster than it can be raised by tectonic loading. The region below 20 km then behaves in a similar way to that above 5 km.

The boundaries at 5 and 20 km are zones of change rather than clear boundaries and are influenced by a range of factors. However, whereas the upper boundary is mainly influenced by overburden pressure (the depth below the free surface) and will change position fast as the fault moves and erosion and deposition occurs. The position of the lower boundary is tied to the very slow migration of geotherms. These effects must be considered in more detailed modelling.

Earthquake rupture initiates in the stressed zone and propagates laterally and also upwards and downwards. Although the driving stress for rupture is confined to the nominal 5–20 km interval, rupture extends outside this zone. The rupture is driven through zones of low stress by first straining the rocks to failure and then fracturing them. As a consequence of passing through an unstressed medium the rupture front loses amplitude as it propagates and leaves large stresses and strains in the surrounding rocks (Fig. 2b). It is the strain associated with this process that is an increment of earthquake folding. Following an earthquake, the stresses relax and the strains become 'frozen' into the rock (Fig. 2c). The tectonic loading gradually restores the stress conditions for the next earthquake. The new event then adds another increment of strain in the low stress zones and it is the cumulative effect of this, plus the cumulative effect of stress relaxation and the effects of loading due to sediment redistribution, that produces the total deformation that forms folds.

The part of the deformation associated with the earthquake itself is rather easy to model. The process is elastic-brittle in

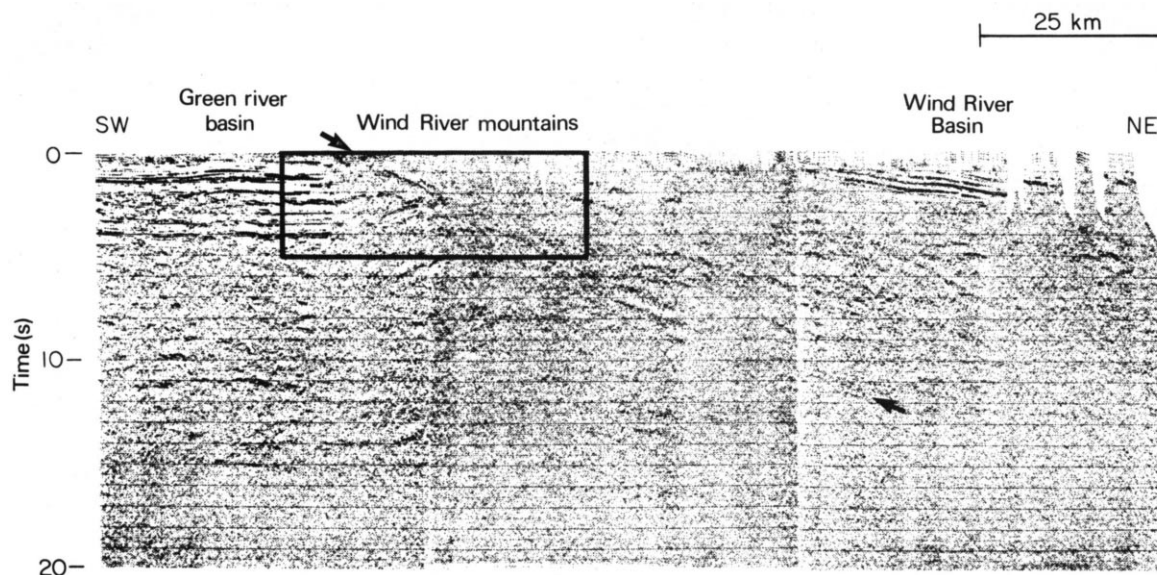


Fig. 1 COCORP data recorded across the Wind River uplift (see discussion in ref. 1). Data are unmigrated. Figure 4a is a migrated interpretation of the area outlined by the rectangle. The Wind River thrust (marked by arrows) underlies the Wind River mountains and dips to the north-east deep into the crust. The Moho in this region probably lies between 14 and 15 s. To convert travel time to approximate depth in kilometres, multiply seconds by 3.

nature and we can use the results of Mansinha and Smylie¹¹ to calculate its relation to fault motion. They derive analytical expressions for the deformation due to a rectangular fault in an elastic half space. By superimposing the deformation due to appropriate distributions of slip on a system of such faults, the deformation due to any fault geometry with any slip distribution may be calculated.

The part of the deformation that is associated with creep relaxation is less easy to calculate. King and Vita-Finzi⁶ comment that the deformation due to stress relaxation below 20 km will always be long wavelength in form and not affect the form of shorter wavelength features. Figure 3 illustrates this by showing the change in deformation of a typical fold when all stress is relaxed below 18 km. Results published by Thatcher and Rundle^{12,13} to describe post-seismic geodetic data, can be interpreted in the same way. The affect of sediment redistribution and erosion are difficult to assess but again may be expected to produce predominantly long wavelength deformation.

Within these limitations we are left with the conclusion that a purely elastic model reasonably describes each increment of folding. Furthermore, because elastic models superpose, a knowledge of the fault orientation and slip distribution of each increment of motion allows us to calculate the total resulting deformation. The reverse is not true as a final deformation distribution can be produced by an infinite number of slip distributions. In these conditions we inevitably make assumptions about the sequence of events that has led to what we observe. An analogous problem is that of finite strain where any number of different histories can produce the same final result¹⁴. The problem is not a very serious one, but must be considered when assessing the significance of our models.

Modelling Wind River thrust

With the foregoing constraints from the mechanics of similar active structures in mind, we now model features of the Wind River thrust profile. The fault observed in the COCORP profiles extends to a depth of 24 km and the best available representation of the footwall structures is Fig. 17 of ref. 4 (our Fig. 4a). The amplitude of the fold in the basement and lower sediments is 0.2 km but this value is uncertain because wavepaths between high velocity Precambrian basement and lower velocity sedimentary rocks are not known. Fold amplitudes could be as much as

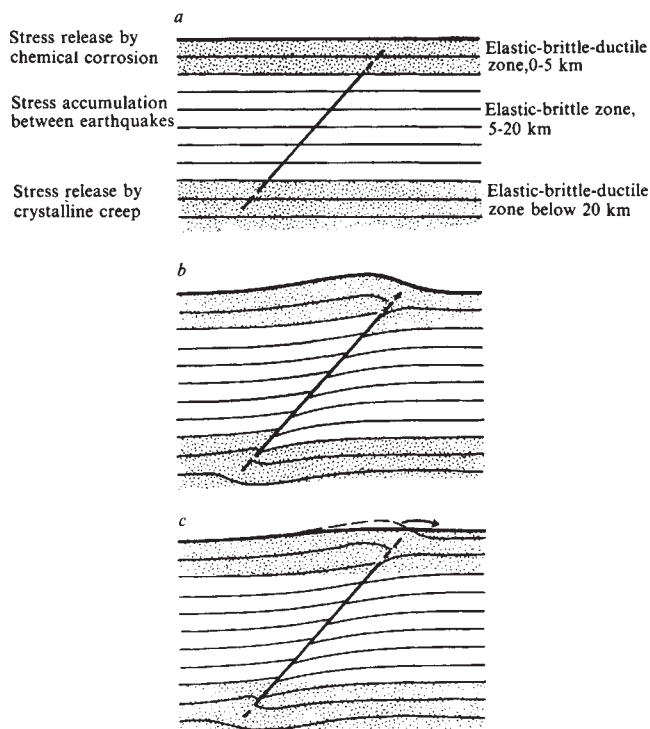


Fig. 2 The stress environment of a shallow continental earthquake. *a*, The conditions before an earthquake; *b*, conditions immediately following; *c*, period of erosion and post seismic stress relaxation.

0.5 km (ref. 4), and perhaps more, as indicated by dotted sedimentary horizons in Fig. 4a. To imitate the form of these structures many models were tried but the basic procedure is indicated by the three deformation plots in Fig. 4b-d. To limit the range of possibilities we model the observed features on the assumption that they result from motion on the observed thrust fault alone. We ignore the possibility that other (seismically transparent) fault structures or decoupling horizons exist. The fact that our model fits the observations without allowing for

the effect of such structures does not mean that other models which include them would not also fit the observations.

The observed fault is clear from the surface to a depth of 24 km. We note (Fig. 4b) that uniform motion on this plane produces none of the observed features of Fig. 4a except for a slight back tilt in the hanging wall. This dip, however, is very much less than that observed in the hanging wall sediments of the Wind River basin. In our modelling procedure we can therefore add a uniform slip over the entire fault plane to any slip distribution that produces the observed features in order to adjust the total fault offset to equal the observed offset.

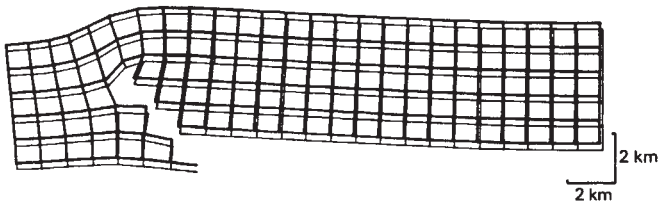


Fig. 3 The near surface deformation associated with a thrust fault with a dip of 45° extending from 3 to 15 km depth. (Each square has a 1-km side.) The heavy lines are for deformation in an elastic half-space and the lighter lines indicate the consequences of relieving all stress below 18 km. The observable effect is a long wavelength relaxation with an amplitude of $<20\%$ of the deformation in a half space. The geologically observable form of the structures is unaltered by the relaxation of stress at depth. The calculations were carried out by Devendra Mishra and are for a fault of infinite extent along strike.

In the second deformation model (Fig. 4c) only one-third of the fault displacement is allowed to enter the upper 3 km and this results in substantial footwall folding. Indeed the form of the footwall feature seems to be largely due to diminished near surface fault slip. However, the northeastward pull-down of the fold near the fault is not as pronounced in this model as it is in Zawislak and Smithson's representations. No modification of the near surface slip distribution was found which would model this asymmetry. To reproduce the pull-down it is necessary to reduce the fault displacement at depth. Reducing the displacement to zero below 18 km (Fig. 4d) goes some way towards modelling the Zawislak and Smithson structure and models the equally plausible dotted alternative (Fig. 4a) fairly well. It also reproduces the dips observed in the hanging wall beds near to the Wind River.

The models described (Fig. 4b–d) all have the same maximum vertical displacement of 4.5 km which is half of that actually inferred for the complete history of Wind River thrust movement. However, we can sum the effects of the models as we choose and thus the final deformation can be explained as being due to the sum of the first (Fig. 4b) and third (Fig. 4d) deformation models. This summing can be done in an infinite number of ways bounded by two possibilities. There could be an initial period of uniform motion followed by a period of non-uniform motion, or the reverse. Between these bounds any mixture is possible. (Although different histories of movement give slightly different results, particularly in the region corresponding to the dipping sedimentary rocks in the Wind River basin.)

However, all models have the feature that the total fault slip has been greater in the 3–18 km depth range than outside this range. Bearing in mind that the 3–18 km depths are not very

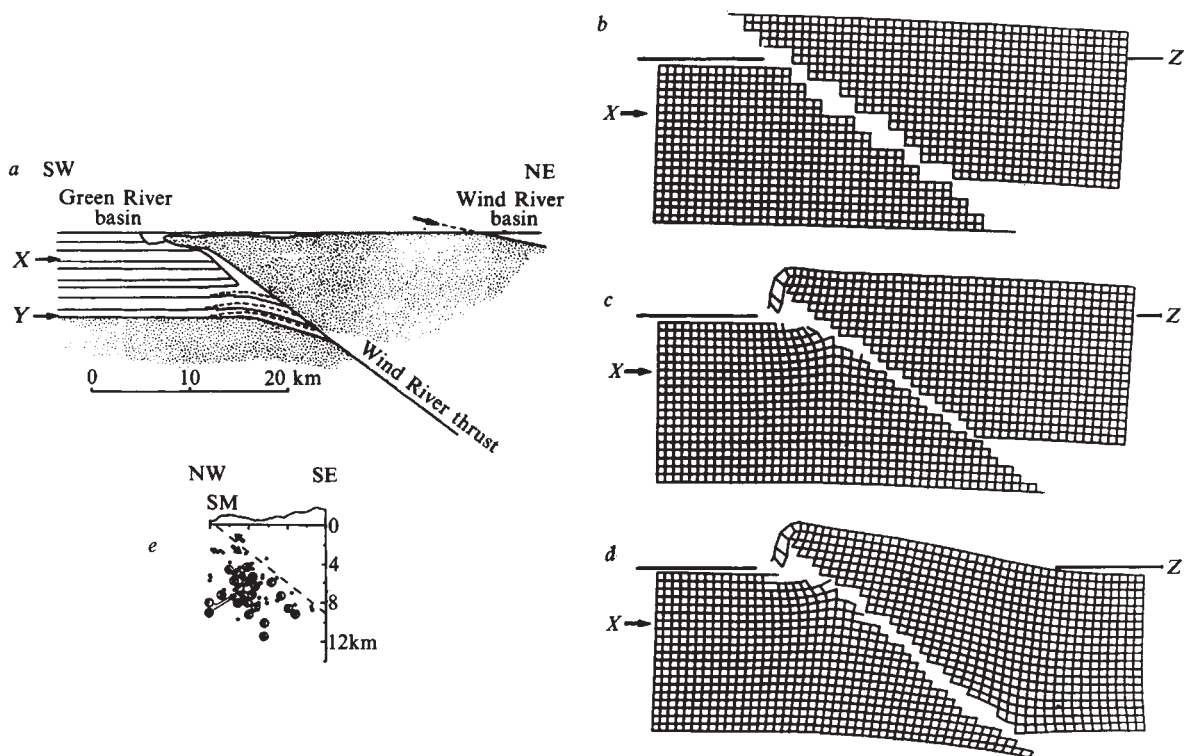


Fig. 4 *a*, Interpretation of the reflection data by Zawislak and Smithson⁴. The basement rocks are indicated by hatching and dotted lines show an alternative interpretation of the footwall deformation, taking into account ambiguities in the wavepaths discussed by Zawislak and Smithson⁴. The scale applies to all parts of the figure. *X*, The base of the sediments under the Green River and Wind River basins. *Y*, Approximate position of ground surface in the Green River basin before thrusting started. *b*, The deformation associated with a fault dipping at 30° with a vertical component of displacement of 4.5 km extending to a depth of 24 km. *Z*, The original top of unfractured crust. *c*, A model with similar geometry except that only one-third of the fault displacement enters the top 3 km of the fault. This causes a pull-up fold in the footwall. *d*, A model that retains the same features as *c* but the fault slip is reduced to zero below 18 km. This model shows a flattening of the dip in the hanging wall well behind the fault. Although this flattening of the dip is a real effect due to the reduction of slip at depth, the actual position is sensitive to the depth at which fault slip diminishes and to small changes of fault dip at this depth. *e*, Aftershocks of the El Asnam (Algeria) earthquake of 10 October 1980. The thrusting aftershocks occur in the same region as the folding seen in the footwall of the Wind River thrust. SM, Sera el Maarouf anticline (topography is plotted at a vertical exaggeration of $3\times$ depth). After Fig. 13b, ref. 15.

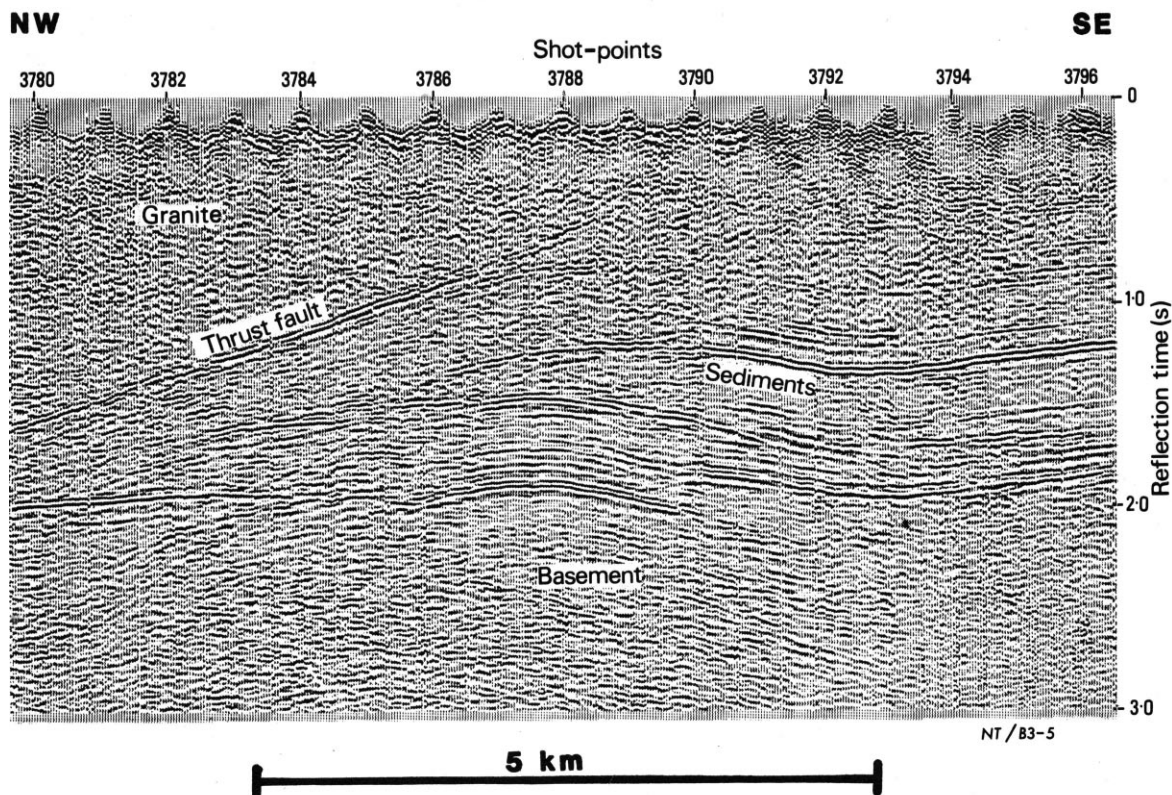


Fig. 5 Seismic reflection section across the northern margin of the Ngalia Basin, Australia, showing overthrusting of granitic rocks onto basinal sediments (from ref. 16). The overthrust fault shows a vertical displacement of ~5 km and a horizontal displacement of ~11.5 km (ref. 17). Reflections from the sedimentary horizons are not displaced upwards in time as they pass under the high-velocity granitic rocks of the overthrust. Therefore some structural downwarping must occur as under the Wind River thrust. Figure supplied by S. P. Mathur.

well defined by the data, this agrees with the nominal 5 and 20 km used for the El Asnam model.

We note that the pull-down effect on the fault is produced by limited fault slip at depth. Should we wish to reproduce the Zawislak and Smithson footwall structure more closely we must reduce the component of slip extending below 18 km accordingly.

Conclusions

The observed folding near the Wind River thrust can be explained if the fault slip on the thrust has been greater over a limited depth range than outside it. This range coincides with the brittle-elastic depth range in which earthquake rupture initiates and propagates in continental earthquakes. At least half of the total fault slip has not extended below 18 km and one-third has not reached the surface. The clearest feature produced by non-uniform slip is a footwall fold. This is a short wavelength feature and our modelling of it is robust and depends on part of the slip on the fault failing to reach the surface. The asymmetry of this fold and back tilting of beds in the fault hanging wall are longer in wavelength and result from diminished slip at depth. These are more likely to be modified by weaknesses in our assumptions about the effect of creep, sediment loading or the presence of other structures. In so far as

our assumptions are valid, our modelling describes the structures in a kinematic fashion which is independent of rock properties.

The fold associated with the Wind River thrust is not unique, a similar fold is observed on seismic reflection data recorded across a granite overthrust at the northern margin of the Ngalia Basin in Australia¹⁵, Fig. 5.

Studies of deformation associated with the El Asnam earthquake suggested the modelling system we use here, but in the El Asnam region only surface folding is visible because no reflection profiles exist. However, the aftershock activity associated with the main part of that thrust¹⁶, Fig. 4e, is in the same part of the footwall that exhibits folding under the Wind River thrust. We therefore conclude that the aftershocks are associated with the active formation of a fold.

If this is the case, we are apparently observing the seismicity associated with the formation of a fold in a brittle environment. This is a surprising conclusion and more seismic reflection data are needed across other basement uplifts to establish whether such footwall folding is a common phenomenon.

This is Earth Sciences contribution no. 391. We thank Devendra Mishra for providing Fig. 3, K. Aki for support at MIT (for G.K.). The work on which this study has been based has been supported by NERC (GR3/3904) and the Royal Society. J.A.B. acknowledges an NERC postdoctoral fellowship.

Received 20 January; accepted 26 July 1983.

1. Smithson, S. B., Brewer, J. A., Kaufman, S., Oliver, J. E. & Hurich, C. A. *J. geophys. Res.* **84**, 5955-5972 (1979).
2. Brewer, J. A., Smithson, S. B., Oliver, J. E., Kaufman, S. & Brown, L. D. *Tectonophysics* **62**, 165-189 (1980).
3. King, P. B. *The Evolution of North America* (Princeton University Press, 1976).
4. Zawislak, R. L. & Smithson, S. B. *Geophysics* **46**, 1684-1701 (1981).
5. King, G., Soufleris, C. & Berberian, M. *Disasters* **5**, 36-46 (1981).
6. King, G. C. P. & Vita-Finzi, C. *Nature* **292**, 22-26 (1981).
7. Yielding, G. et al. *Earth planet. Sci. Lett.* **56**, 287-304 (1981).

8. Ruegg, J. C., Kasser, M., Tarantola, A., Lepine, J. C. & Chouikrat, B. *Bull. Seism. Soc. Amer.* **72**, 2227-2244 (1982).
9. Thommeret, Y., Vita-Finzi, C. & King, G. C. P. *Earth planet. Sci. Lett.* **63**, 137-138 (1983).
10. Sibson, R. H. *Bull. seism. Soc. Am.* **72**, 151-163 (1982).
11. Mansinha, L. & Smylie, D. E. *Bull. seism. Soc. Am.* **61**, 1433-1440 (1971).
12. Rundle, J. B. *J. geophys. Res.* **83**, 5937-5945 (1978).
13. Rundle, J. B. & Thatcher, W. *Bull. seism. Soc. Am.* **70**, 1869-1886 (1980).
14. Ramsay, J. G. *Folding and Fracturing of Rocks* (McGraw-Hill, New York, 1967).
15. Mathur, S. P. *BMR J. Austr. Geol. Geophys.* **1**, 277-286 (1976).
16. Wells, A. T., Moss, F. J. & Sabitay, A. *Austr. Petrol. Explor. J.* **12**, 144-151 (1972).
17. Ouyed, M., Yielding, G., Hatzfeld, D. & King, G. C. P. *Geophys. J. R. astr. Soc.* **73**, 605-639 (1983).

Cloned human phenylalanine hydroxylase gene allows prenatal diagnosis and carrier detection of classical phenylketonuria

Savio L. C. Woo, Alan S. Lidsky, Flemming Güttler*, T. Chandra & Kathryn J. H. Robson

Howard Hughes Medical Institute, Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030, USA

* John F. Kennedy Institutet, DK-2600, Glostrup, Denmark

The human gene for the hepatic enzyme phenylalanine hydroxylase has been cloned and used to analyse the phenylalanine hydroxylase locus in the human genome. The detection of polymorphisms in this locus by several restriction enzymes has allowed feasibility studies of prenatal diagnosis of classical phenylketonuria and identification of carriers of the trait. Results indicate that these services could be provided for up to 75% of all families with phenylketonuric children in the general Caucasian population.

CLASSICAL phenylketonuria (PKU) is a human genetic disorder caused by an inborn error in aromatic amino acid metabolism. Fölling¹ first observed, some 50 years ago, that patients with this condition excrete large quantities of phenylpyruvate which is a minor metabolite of phenylalanine in normal individuals. Subsequently, it was discovered that livers of such patients were unable to convert phenylalanine to tyrosine due to a deficiency of the enzyme phenylalanine hydroxylase (PH)²⁻⁴. The absence of this enzymatic activity causes persistent hyperphenylalaninaemia causing disturbances in tyrosine and tryptophan metabolism⁵, and diminished formation of catecholamines, melanin and serotonin is typical in untreated PKU patients. The clinical manifestation is an irreversible impairment of postnatal brain development resulting in severe mental retardation of the untreated children^{1,6}. PH deficiency, regardless of phenotypic variations, is transmitted as an autosomal recessive trait. Massive screening of over 5 million neonates has shown that the prevalence of classical PKU among Caucasians ranges from 1/5,400 in Ireland to 1/16,000 in Switzerland⁷. The collective frequency in eight Western European countries and the United States is ~1/10,000 so that 2 out of every 100 individuals in the general population are carriers of the PKU trait^{8,9}.

Newborn infants affected with classical PKU can be readily identified by the Guthrie¹⁰ test and subsequently treated with a low phenylalanine diet^{11,12}. The dietary correction, however, must be implemented rigidly over a prolonged period of time to be effective^{13,14}. Termination of treatment in the middle of the first decade of life is apparently followed by a small but significant decline in IQ scores^{15,16}. More recent studies have indicated that there might be subtle changes in cerebral function in children with classical PKU in whom the phenylalanine diet has been discontinued¹⁷, and that there is a definite trend towards later discontinuation of the dietary correction¹⁸. Furthermore, as PH is a hepatic enzyme and not present in serum or fibroblast cells, there is no available methodology for prenatal diagnosis of the hereditary disorder. We have previously reported the purification of rat liver PH mRNA by polysome immunoprecipitation and the cloning of its cDNA¹⁹. We now report the use of the rat cDNA clone to identify the corresponding human cDNA clones, and the use of the latter to analyse classical PKU by restriction fragment length polymorphism in the PH locus.

Cloning of human PH cDNA

Using a previously cloned rat PH cDNA clone (prPH98) as a specific hybridization probe, a human liver cDNA library²⁰

comprising 40,000 independent transformants was screened. Recombinant plasmid DNAs were isolated from the bacterial colonies that yielded strong hybridization signals. The DNA samples were analysed by Southern hybridization after digestion with the restriction enzyme *HhaI* which cleaved the plasmid vector DNA into small fragments, but did not cut the inserted human DNA in any of the clones (K.J.H.R., unpublished results). As positive controls of the experiment, two previously cloned rat PH cDNA recombinants¹⁹ were included and strong hybridization signals were obtained (Fig. 1A, lanes 1, 2). The specificity of the hybridization probe has been demonstrated by the lack of hybridization with the vector DNA pBR322 (Fig. 1A, lane 3) and with a human α_1 -antitrypsin cDNA clone (Fig. 1A, lane 4). Specific hybridization signals were obtained for the PH cDNA clones identified from the human liver cDNA library (Fig. 1A, lanes 5-9). The lengths of the inserted human DNA in these recombinant plasmids ranged from 0.3 to 1.4 kilobases (kb). Partial DNA sequence analysis has indicated that the extent of sequence homology between the rat and human PH cDNA clones is ~90% (K.J.H.R. *et al.*, unpublished results). The two recombinants containing the largest inserted human DNA fragments (phPH72 and phPH73) were used as hybridization probes for the subsequent experiments.

PH gene in classical phenylketonuria

Genomic DNAs were isolated from two PKU cell lines obtained from the Human Mutant Cell Repository (GM937 and GM2406) and from two normal individuals. The DNA preparations were digested with various restriction enzymes, followed by Southern hybridization using the human PH cDNA clones as specific probes. Identical hybridization signals were obtained from all four DNA preparations after digestion with more than 10 enzymes; results for three of these are shown in Fig. 1B. The results indicate that the PH gene is not only still present in the genome of cells derived from PKU patients, but that the overall organization of the gene also remains unchanged. Comparison of densitometer tracings of the gel lanes containing normal and PKU DNAs has shown that the hybridization signals generated by the PKU DNA samples are not the result of compound heterozygotes having deletions in different regions of the PH gene in the two chromosomes present in each cell line. As the cDNA probes were only 1.4 kb long and the human PH mRNA is comprised of ~2,500 nucleotides, it is possible that there could be a deletion or rearrangement in regions of the PH gene that are not represented in the cDNA clones. Thus, we can only conclude that classical PKU, at least in these

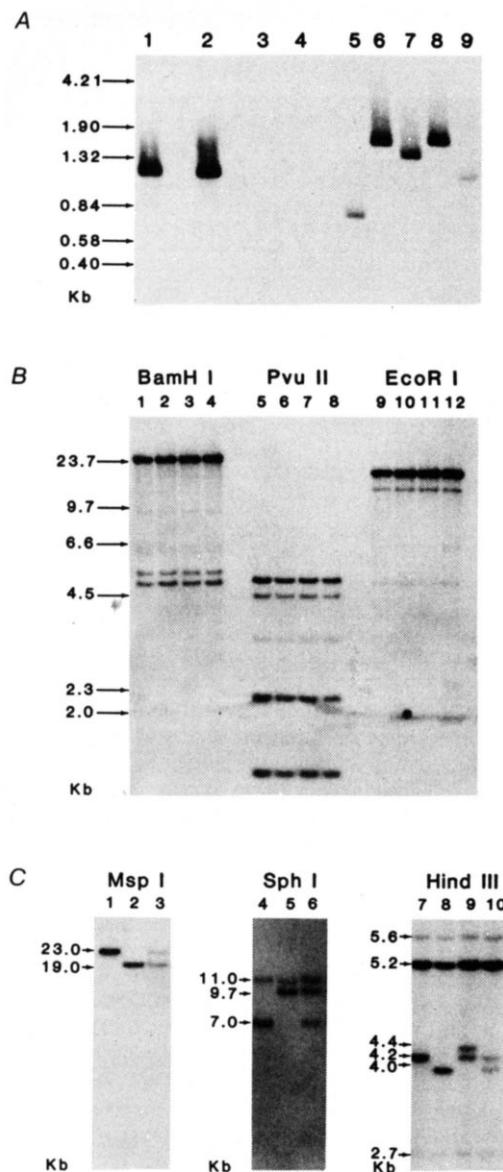


Fig. 1 A, Analysis of human PH cDNA clones by cross-hybridization using as a probe the inserted *Pst*I DNA fragment isolated from prPH98, a previously reported rat PH cDNA clone¹⁹. Lane 1, prPH91; lane 2, prPH98; lane 3, pBR322; lane 4, a human α_1 -antitrypsin cDNA clone; lanes 5-9, five human PH cDNA clones (phPH71-phPH75, respectively). B, Southern blot analysis of genomic DNA samples isolated from lymphocytes of normal individuals (lanes 1, 2, 5, 6, 9 and 10), and PKU fibroblast cell lines GM0937 (lanes 3, 7 and 11) and GM2406 (lanes 4, 8 and 12) obtained from the Human Mutant Cell Depository. The hybridization probes used were labelled *Pst*I fragments isolated from phPH72 and phPH 73 DNA. C, Southern blot analysis of 10 μ g of genomic DNAs isolated from normal Caucasians after digestion with *Msp*I (lanes 1-3), *Sph*I (lanes 4-6) and *Hind*III (lanes 7-10). The Caucasians were chosen from a randomly selected panel of 20 individuals on the basis of their representative polymorphic patterns. Southern hybridization was carried out as described below for B. Other enzymes used for analysis that did not yield polymorphic patterns were: *Bam*HI, *Bgl*II, *Ava*II, *Eco*RI, *Taq*I, *Bcl*I, *Kpn*I, *Pst*I, *Xho*I, *Sst*I, *Stu*I and *Pvu*II.

Methods: prPH98 DNA was digested with *Pst*I and the inserted rat DNA sequence was isolated by preparative polyacrylamide gel electrophoresis (PAGE). Any contaminating plasmid DNA was removed from the isolated *Pst*I fragment by digestion with *Hha*I and subsequent purification by preparative PAGE. The rat PH cDNA probe was then labelled by nick-translation²⁵ to a specific radioactivity of 2×10^8 c.p.m. μ g⁻¹ and used to screen a human liver cDNA library¹⁸ comprised of 40,000 independent transformants. Five clones were identified and DNAs isolated from them were analysed by agarose gel electrophoresis followed by Southern blot analysis²⁶. For various DNA samples, 1 μ g was digested to completion with the enzyme *Hha*I before electrophoresis. B and C, DNA was prepared from isolated lymphocytes and cultured cells by proteinase K (100 μ g ml⁻¹) and SDS (0.5%) treatment for 16 h at 55 °C, followed by phenol-chloroform-isoamylalcohol (25:24:1) extraction and precipitation with ethanol. DNA was redissolved in TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 7.0) and treated with 50 μ g ml⁻¹ of DNase-free ribonuclease at 37 °C for 1 h, followed by phenol extraction and alcohol precipitation. The DNA pellets were again redissolved in TE buffer; 10 μ g of DNA were then digested with restriction endonucleases at 2 units per μ g DNA according to manufacturer's specifications, followed by electrophoresis in 1% agarose gels. DNA was transferred from the agarose gels to nitrocellulose filters according to the method of Southern²⁶ and baked at 68 °C for 3 h. The filters were then pretreated and hybridized to labelled phPH72 and phPH73 DNAs in the presence of 10% dextran sulphate according to the procedure of Wahl *et al.*²⁶. Filters were then washed and exposed to Fuji Medical X-ray film at -70 °C with intensifying screens.

two cases, is not caused by deletion of the entire *PH* gene. This observation has subsequently been extended to include several additional phenylketonuric individuals analysed in the following sections.

Polymorphism in the human *PH* locus

There are many nucleotide substitutions in the genomes of different individuals²¹ that are not expressed phenotypically, but which can sometimes be detected by restriction enzymes. Such restriction site polymorphisms result in a difference of the DNA fragment length, which can be readily detected by agarose gel electrophoresis followed by Southern hybridization using specific hybridization probes. Restriction fragment length polymorphism has since become a powerful tool for molecular analysis of human genetic disorders.

Genomic DNAs isolated from 20 random Caucasians in the Houston metropolitan area were analysed using a battery of restriction enzymes, and three enzymes that yielded polymorphic patterns in the *PH* locus were identified. The enzyme *Msp*I generated a single hybridization band at 23 kb in some individuals and at 19 kb in others (Fig. 1C, lanes 1 and 2) indicating the existence of two genotypic alleles. Indeed, heterozygotes containing both bands have also been detected (Fig. 1C, lane 3). The frequencies of the alleles associated with the 23-kb and 19-kb bands are 0.444 and 0.556, respectively.

Furthermore, as there is only a single hybridization band in genomic DNAs of normal homozygotes and since the chromosomal gene appears to be greater than 20 kb (A. G. DiLella, unpublished results), the *PH* gene is probably a unique sequence in the human genome with no apparent pseudogenes. Consequently, all hybridization signals generated by digestion of genomic DNAs with other enzymes must also be derived from the *PH* locus itself, but not from other chromosomal regions containing homologous sequences.

The existence of a second allelic system in the human *PH* locus has been identified by the enzyme *Sph*I, which cleaved the DNAs into either a 7.0-kb or a 9.7-kb fragment in different individuals (Fig. 1C, lanes 4, 5). The presence of both bands in heterozygotes has also been detected (Fig. 1C, lane 6). The 11.0-kb band present in every individual is a neighbouring DNA fragment and is non-polymorphic. In this case, the frequency for the 9.7-kb and 7.0-kb alleles is 0.059 and 0.941, respectively.

The enzyme *Hind*III revealed a three-allelic system, in which most individuals are either homozygous or heterozygous for the 4.2-kb and 4.0-kb bands (Fig. 1C, lanes 7, 8 and 10), the frequencies of which are 0.706 and 0.235, respectively. A rarer 4.4-kb band has also been detected in some individuals, with a frequency of 0.059 (Fig. 1C, lane 9). The existence of these allelic systems permits the analysis of classical PKU by restriction fragment length polymorphism.

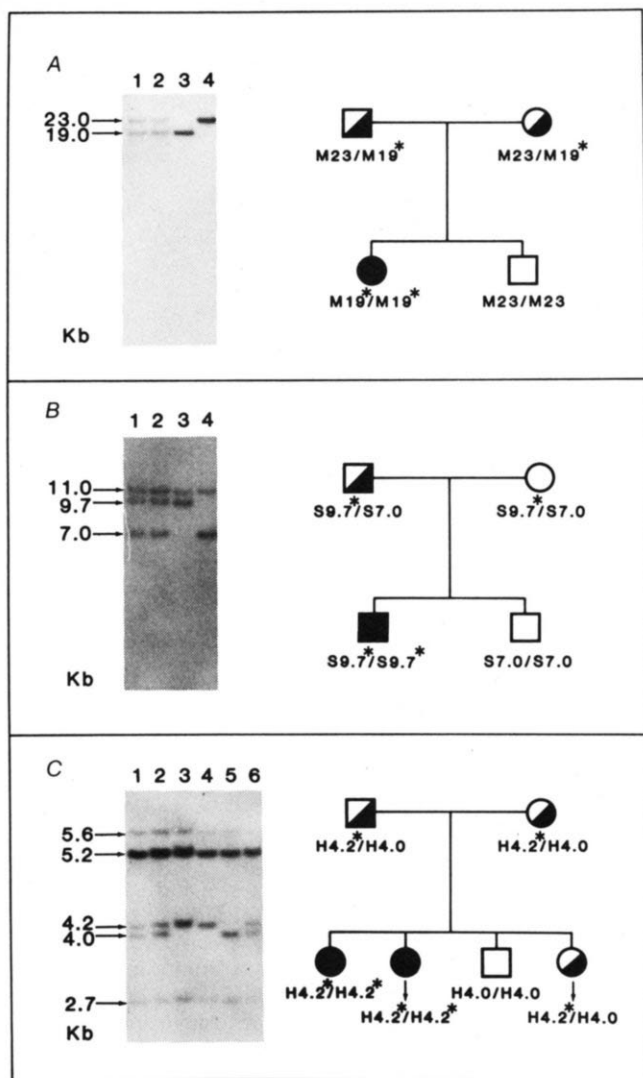


Fig. 2 Analysis of three Danish PKU families using *MspI* (A), *SphI* (B) and *HindIII* (C) by restriction fragment length polymorphisms in the *PH* locus. DNA was isolated from lymphocytes of the family members and subjected to Southern blot analysis using phPH72 and phPH73 DNA as probes (see Fig. 1B). In A and B: lanes 1, father; lanes 2, mother; lanes 3, proband; lanes 4, unaffected sibling with serum phenylalanine concentration [Phe] of 17 and 13 mg l⁻¹ in A and B, respectively. C: Lane 1, father; lane 2, mother; lanes 3 and 4, two affected children; lanes 5 and 6, two unaffected children with serum [Phe] of 42 and 35 mg l⁻¹, respectively. The segregation of the PKU alleles (*) with the appropriate polymorphic bands is shown schematically on the right of each autoradiograph. All PKU proband in these families have phenylalanine tolerance levels <0.13 mmol per kg per day, and all parents have discriminant values below 6 (ref. 28). Phenylalanine loading was not performed on the unaffected siblings who were all negative to the Guthrie test¹⁰ for hyperphenylalaninaemia.

Kindred analysis of classical phenylketonuria

Seven Danish PKU families having one or two affected children were analysed by restriction fragment length polymorphism. In one such family, both parents are heterozygous for the 23-kb and 19-kb bands as revealed by the enzyme *MspI* (Fig. 2A, lanes 1, 2). As both parents are obligatory carriers of the PKU trait, each must contain a mutant *PH* gene which could be associated with either allele. The proband is homozygous for the 19-kb band (Fig. 2A, lane 3), indicating that the mutant *PH* genes of both parents reside in the 19-kb alleles. An unaffected sibling in this family is homozygous for the 23-kb band (Fig. 2A, lane 4), indicating inheritance of the normal *PH* genes from both parents, and should therefore be free of the phenylk-

etonuric trait. A diagram illustrating the mendelian segregation of the two alleles in this family is also shown (Fig. 2A). These results indicate that prenatal diagnosis for classical PKU will be possible in future pregnancies of this family by analysis of DNA isolated from amniotic cells of fetuses using the *MspI* restriction fragment length polymorphism test. A fetus homozygous for the 19-kb band will be phenylketonuric, one that is heterozygous for the two *MspI* bands will be a carrier, and one that is homozygous for the 23-kb band will be free of the PKU trait.

The mutant *PH* genes in some PKU families can be identified by *SphI* and *HindIII* polymorphisms. In one such family, both parents are heterozygous for the 9.7-kb and 7.0-kb *SphI* bands (Fig. 2B, lanes 1, 2). The proband is homozygous for the 9.7-kb band (Fig. 2B, lane 3), while an unaffected sibling is homozygous for the 7.0-kb band and is therefore not a carrier of the PKU trait (Fig. 2B, lane 4). In another family, *HindIII* polymorphism has been detected in both parents who are heterozygous for the 4.2-kb and 4.0-kb bands (Fig. 2C, lanes 1, 2). The proband is homozygous for the 4.2-kb band (Fig. 2C, lane 3), indicating that the mutant *PH* genes in this family are associated with the 4.2-kb bands. A second affected child in this family is also homozygous for the 4.2-kb band (Fig. 2C, lane 4), indicating that the segregation of the PKU alleles and disease state is concordant. This interpretation is further supported by the observation that two unaffected siblings in the same family have inherited different *PH* alleles from the parents. One sibling is homozygous for the 4.0-kb bands (Fig. 2C, lane 5) and is free of the trait, another sibling is heterozygous for the 4.2-kb and 4.0-kb bands (Fig. 2C, lane 6) and must be a trait carrier. The observation that allelic segregation between the probands and unaffected siblings is discordant in all families analysed, strongly suggests that classical PKU is caused by an underlying mutation in the *PH* gene itself and not by a transregulatory mechanism. These results are in agreement with previous observations that *PH* from patients with classical PKU appears to be a structurally altered enzyme having <1% of the normal activity²².

Haplotype analysis

Haplotype identification of genetic alleles in man using a combination of restriction enzymes is an extremely powerful tool for analysis of human genetic disorders. The exploitation of this technology in the analysis of various forms of thalassaemia has recently been well documented²³. The application of haplotype analysis of the *PH* gene in one of the Danish families is shown in Fig. 3. In this family, both the mother and the proband are homozygous for the 7.0-kb *SphI* band while the father is heterozygous (Fig. 3B, lanes 1–3). Thus, the father's mutant *PH* gene must be associated with the 7.0-kb allele which has also been transmitted into an unaffected sibling homozygous for the 7.0-kb band (Fig. 3B, lane 5). As this sibling is also homozygous for the 4.0-kb *HindIII* band (Fig. 3A, lane 5), the paternal mutant *PH* gene must have a haplotype of H4.0:S7.0, where H and S represent *HindIII* and *SphI*, respectively. The fact that both parents and the proband are heterozygous for the 4.2-kb and 4.0-kb *HindIII* bands (Fig. 3A, lanes 1–3) suggests that the maternal mutant gene is associated with the 4.2-kb allele and has a haplotype of H4.2:S7.0. The haplotypes of the normal *PH* genes in the mother and the father must therefore be H4.0:S7.0 and H4.2:S9.7, respectively.

Once the haplotypes of all four *PH* genes in the parents are established, it becomes apparent that the proband inherited the mutant genes from both parents, while the unaffected sibling of the H4.0:S7.0 haplotype has inherited the paternal mutant *PH* gene plus the maternal normal *PH* gene, and is thus a carrier of the PKU trait (Fig. 3C). A second unaffected sibling has inherited the normal *PH* genes from both parents and is consequently free of the trait (Fig. 3C). Thus, families that are non-informative by genotyping using a single enzyme could become informative by haplotyping using a combination of restriction enzymes.

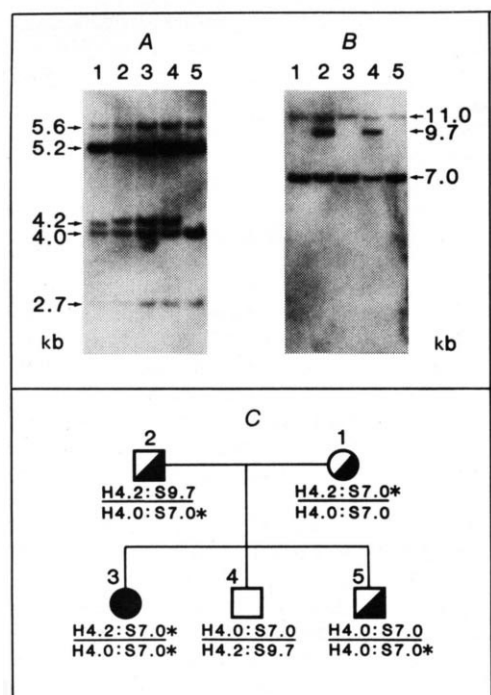


Fig. 3 Haplotype analysis of the *PH* locus in a Danish PKU family. DNA samples (10 μ g) were digested with either *Hind*III (A) or *Sph*I (B) and applied to each lane in the gel as follows: lane 1, mother; lane 2, father; lane 3, proband; lane 4, an unaffected sibling (serum [Phe] = 25 mg l^{-1}); lane 5, a second unaffected sibling (serum [Phe] = 22 mg l^{-1}). Southern blotting was performed as described in Fig. 1 legend, and the criteria for the PKU proband were the same as in Fig. 2. C, Diagram showing the haplotypes of the *PH* gene for each family member and the segregation of PKU alleles (*) in this family.

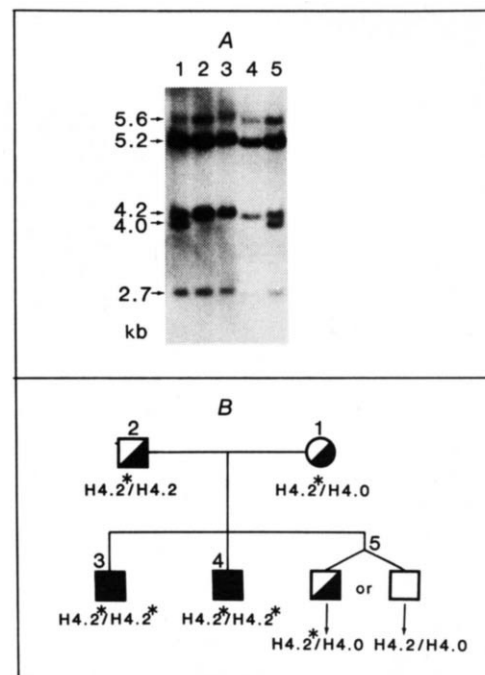


Fig. 4 Exclusion of recessive homozygotes in a PKU family by *Hind*III polymorphism analysis (A) and the corresponding mendelian segregation of the PKU alleles (B). Lane 1, DNA from mother; lane 2, DNA from father; lanes 3 and 4, DNAs of two affected children; and lane 5, an unaffected sibling's DNA (serum [Phe] = 25 mg l^{-1}).

Exclusion of phenylketonuric homozygotes

In some families, only one parent is a haplotype heterozygote. The other parent, being a haplotype homozygote, is non-informative for this analysis. In one such family, the mother is heterozygous for the 4.2-kb and 4.0-kb *Hind*III bands, while the father is homozygous for the 4.2-kb band (Fig. 4A, lanes 1, 2). There are two affected children in this family, and both are homozygous for the 4.2-kb band (Fig. 4A, lanes 3, 4), indicating that the maternal mutant *PH* gene is associated with the 4.2-kb allele. As the unaffected sibling is heterozygous for the two bands (Fig. 4A, lane 5), the possibility of this individual being of the PKU phenotype can definitely be excluded on the basis that the maternal normal *PH* gene (the 4.0-kb allele) has been transmitted to this individual (Fig. 4B). Any siblings having the homozygous 4.2-kb genotype must have inherited the maternal mutant *PH* gene plus either the paternal normal or mutant *PH* gene, and would consequently either be trait carriers or phenylketonuric, respectively. In a hypothetical future pregnancy within this family, the probability for exclusion of the fetus being a PKU homozygote will be 50%, which is the probability of transmission of the paternal normal *PH* gene. This percentage for exclusion by restriction fragment length polymorphism analysis will also be true for all families having a homozygous parent and a heterozygous parent.

Probability of diagnosis

Given the individual frequencies of the three *Hind*III alleles, the two *Msp*I alleles and the two *Sph*I alleles of the *PH* gene, the per cent heterozygosity and homozygosity among Caucasians has been calculated to be 75% and 25%, respectively. In the general population, the probabilities of mating events between two heterozygotes, two homozygotes, and a heterozygote with a homozygote are governed by the binomial expansion $(75\% + 25\%)^2$. Thus the probability of two heterozygotes mating would

be $(75\%)^2$ or 56.25% and in such cases, if the haplotypes of the mutant *PH* genes of the parents of a PKU child can be established, complete prenatal diagnosis can be expected by analysis of fetal DNA obtained through amniocentesis. If the mating is between two homozygotes, which has a probability of $(25\%)^2$ or 6.25%, there can be no diagnosis or exclusion and obviously amniocentesis should not be performed. Finally, the probability of mating events between a heterozygote and a homozygote is $2 \times (75\%)(25\%)$ or 37.5%. As only 50% of the resulting fetuses can be excluded from being a PKU homozygote, the overall probability of excluding recessive homozygotes will be 18.75%. Thus, genetic service by DNA analysis can be provided to 56.25% + 18.75% = 75% of the PKU families in ideal conditions.

Conclusions

A previously cloned rat *PH* cDNA clone has been used to identify the corresponding human cDNA clones in a human liver cDNA library by cross-hybridization. These human cDNA clones have been used as hybridization probes to analyse classical PKU. We have obtained conclusive evidence for the presence of the *PH* gene in the cellular DNA of the PKU individuals, suggesting that classical PKU is not caused by a deletion of the entire *PH* gene. The possibility of heterozygous deletion in the *PH* gene being the cause of PKU can be excluded in the families analysed, since the PKU alleles in these families behaved in identical manner to the normal genes in a Southern blot analysis.

Since the initial observation by Kan and Dozy²⁴ that there is a *Hpa*I restriction site polymorphism at the 3' flanking region of the human β -globin gene which can be used for diagnosis of sickle-cell anaemia, restriction fragment length polymorphism has developed into an extremely powerful tool to analyse human genetic disorders. There are over 10 enzymes that yield polymorphic patterns from the 60-kb human β -globin locus, and the application of these polymorphisms for prenatal diagnosis of various kinds of β° and β^+ thalassaemia by linkage analysis has been well documented²⁵. By using the cloned human *PH* cDNA probe, we have identified three restriction enzymes that yield polymorphic patterns from the *PH* locus, resulting in a total of

12 theoretical haplotypes of the gene. Based on a rather small sample size (20 random individuals), and not taking into consideration the possibility of linkage disequilibrium between these polymorphic sites, it has been calculated that the per cent heterozygosity within the general Caucasian population is 75%. Obviously, this percentage may be modified in the future when a much larger sample size is analysed. Nevertheless, this high level of heterozygosity suggests that prenatal diagnosis can be performed for 56.25% of the random PKU families and exclusion of recessive homozygotes can be achieved in an additional 18.75% of the families. These probabilities are, however, optimal estimates in that the haplotypes of all four *PH* genes in a particular family may not be identifiable by comparison with the proband alone. A typical example is provided in the family shown in Fig. 3. Without the availability of the informative sibling, complete diagnosis would not have been possible, unless data from other siblings or grandparents allowed the phase of the haplotypes of the *PH* gene in the parents to be

established. In ideal conditions, genetic services can be provided for 75% of the random PKU families in the general Caucasian population at present. These numbers should improve in the future with the identification of additional alleles by restriction polymorphism when more DNA in the human *PH* gene is cloned and used as hybridization probes. Consequently, prenatal diagnosis of classical PKU and carrier detection should become a reality for most of the PKU families in the general population.

We thank Drs Robert Nussbaum, Stephen P. Daiger, John Phillips III and Haig Kazazian for helpful discussions, Dr Arthur Beaudet for critical reading of the manuscript and Ms Sharon Moore for technical assistance. This work was partially supported by Danish MRC grants 512-5313 and 512-6840 and a Danish Health Insurance Foundation grant H11/28/78 to F.G. and by a March of Dimes Birth Defects Foundation grant (I-566) and NIH grant HD-17711 to S.L.C.W., who is also an Investigator of the Howard Hughes Medical Institute.

Received 11 May; accepted 18 August 1983.

- Fölling, A. *Nord. Med.* **8**, 1054-1059 (1934).
- Jervis, G. A. *Proc. Soc. exp. Biol. Med.* **82**, 514-515 (1953).
- Udenfriend, S. & Bessman, S. P. *J. biol. Chem.* **203**, 961-966 (1953).
- Kaufman, S. in *Advances in Neurochemistry* Vol. 2 (eds Agranoff, B. W. & Aprison, M. H.) 1-132 (Plenum, New York, 1976).
- Blau, K. in *Aromatic Amino Acid Hydroxylase and Mental Disease* (ed. Youdim, M. B. H.) 79-139 (Wiley, New York, 1979).
- Fölling, A. *Hoppe-Seyler's Z. physiol. Chem.* **227**, 169-176 (1934).
- Thalhammer, O. *Humangenetik* **30**, 273-286 (1975).
- Bickel, H. *et al. Acta paediat. scand.* **62**, 413-416 (1973).
- Scriver, C. R. & Rosenberg, L. E. in *Amino Acid Metabolism and Its Disorders*, 290-337 (Saunders, Philadelphia, 1973).
- Guthrie, R. & Susi, A. *Pediatrics* **32**, 338-343 (1963).
- Bickel, H., Gerrard, J. & Hickmans, E. M. *Acta paediat. scand.* **43**, 64-67 (1954).
- Armstrong, M. D. & Tyler, F. H. *J. clin. Invest.* **34**, 565-580 (1955).
- Smith, I. *et al. Br. med. J.* **2**, 723-726 (1978).
- Koch, R., Friedman, E. G., Williamson, M. L. & Azen, C. G. *J. inher. Metab. Dis.* **5**, Suppl. 1, 63-64 (1982).
- Scriver, C. R. & Clow, C. L. *New Engl. J. Med.* **303**, 1336 (1980a).
- Scriver, C. R. & Clow, C. L. *New Engl. J. Med.* **303**, 1394 (1980b).
- Koch, R., Azen, C. G., Friedman, E. G. & Williamson, M. L. *J. Pediatr.* **100**, 870-875 (1982).
- Schuetz, V. E., Gurda, R. F. & Brown, E. S. *Am. J. publ. Hlth* **70**, 498-503 (1980).
- Robson, K. J. H., Chandra, T., MacGillivray, R. T. A. & Woo, S. L. C. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4701-4705 (1982).
- Chandra, T., Stackhouse, R. & Woo, S. L. C. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1845-1848 (1983).
- Jeffreys, A. J. *Cell* **18**, 1-10 (1979).
- Friedman, P. O., Fisher, D. B., Kang, E. S. & Kaufman, S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 552-556 (1973).
- Orkin, S. H. *et al. Nature* **296**, 627-631 (1982).
- Kan, Y. W. & Dozy, A. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5631-5635 (1978).
- Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184-1188 (1975).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).
- Güttler, F. *Acta. paediat. scand. Suppl.* **280**, 1-80 (1980).

The terminal nucleotides of retrovirus DNA are required for integration but not virus production

Antonito T. Panganiban & Howard M. Temin

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, Wisconsin 53706, USA

Deletion of specific nucleotides at either end of the long terminal repeat of the avian retrovirus, spleen necrosis virus, results in replication-competent but integration-defective virus. This result supports two conclusions: (1) the 5-base pair terminal inverted repeats and three to seven adjacent nucleotides are required for integration; (2) integration of retrovirus DNA is not required for retrovirus gene expression.

RETROVIRUSES are a family of single-stranded RNA viruses that replicate through a DNA intermediate. The viral DNA is initially found in the cytoplasm of the host cell and subsequently in the nucleus¹⁻³, where integration into host cell DNA occurs, resulting in the formation of the provirus⁴⁻⁶. Although small amounts of circular DNA are found in the nucleus^{1,3}, the predominant form of unintegrated viral DNA is a linear molecule containing two copies of the long terminal repeat (LTR)^{7,8}. These repeats are generated from sequences at both ends of the viral RNA during DNA synthesis⁹⁻¹³. The DNA of the provirus is at least grossly colinear with the linear form of the DNA¹⁴⁻¹⁶.

Nucleotide sequence characterization of the termini of retrovirus DNA has revealed several important features. Examination of the outside ends of unintegrated DNA¹⁷ and the inside

termini of the LTR from cloned integrated DNA¹⁸⁻²² indicates that the ends of the LTRs from unintegrated linear DNA contain short inverted repeats ranging from 5 to 13 base pairs (bp) depending on the individual retrovirus species. Furthermore, the process of integration has two consequences for the structure of the viral DNA and the host DNA at the site of insertion: the terminal 2 bp from each inverted repeat are lost from the viral DNA^{21,23-28} and a short, precise, direct duplication of the cellular DNA results^{21,23,25,27,28}. The fact that establishment of the provirus takes place in such an exact manner and that alteration of the terminal region of the viral DNA occurs indicates that the ends of the viral DNA may be important for recognition by the enzyme(s) required for DNA integration.

We have examined the role of the terminal region of the viral

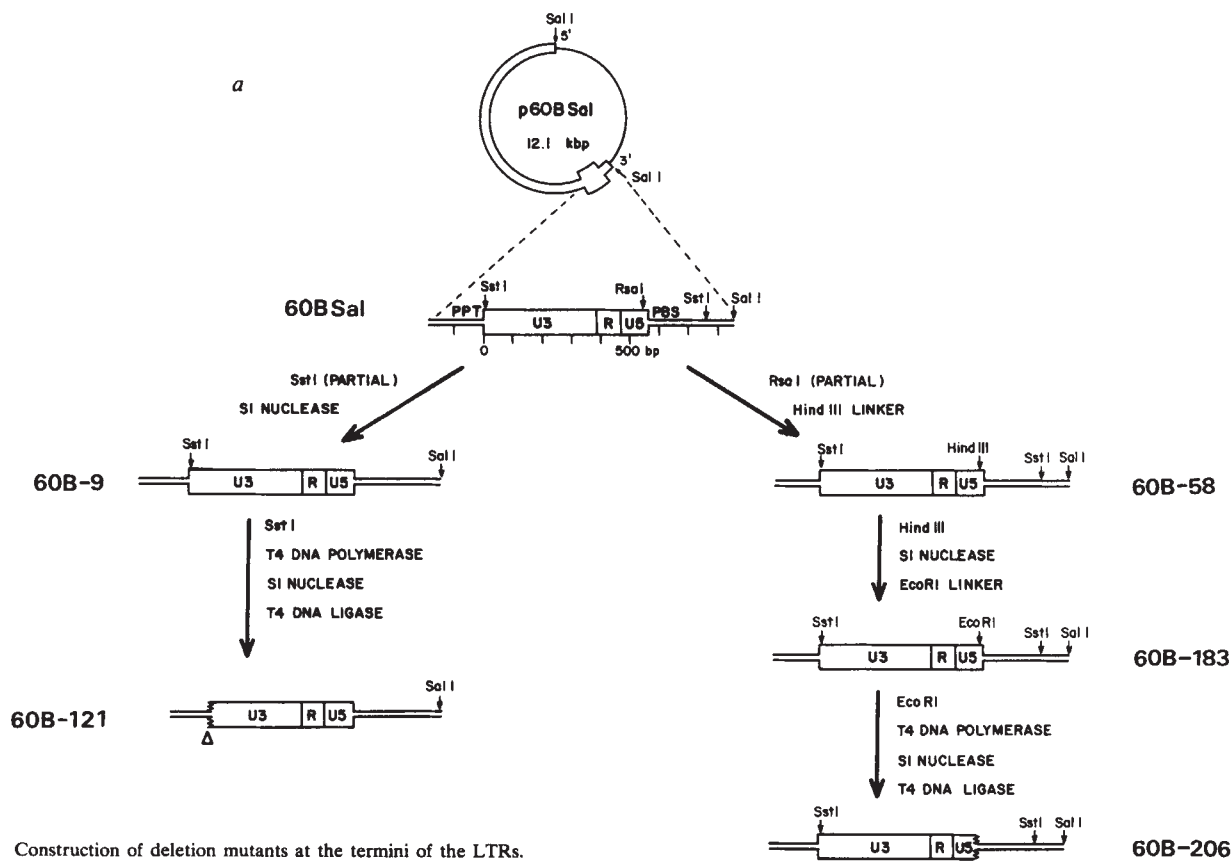


Fig. 1 Construction of deletion mutants at the termini of the LTRs. **a**, Strategy used to generate deletion mutants. p60BSal is shown at the top^{31,32}. The open bar represents the SNV insert and the wide bar designates the LTR; the line represents pBR322 sequences. For simplicity, the remainder of the figure depicts only the altered region of the DNA containing the LTR. p60B-121 and p60B-206 are representative of a group of plasmids bearing short deletions: p60B-121, -144, -150, -147 and -128 contain deletions at the end of U3 while p60B-206, -211, -207, -205 and -227 contain deletions at the end of U5. Triangles show the regions deleted from the termini of the LTR. PPT and PBS indicate the polyurine tract and primer binding site adjacent to the LTR, respectively. **b** Depicts the concatemer formed by digestion of p60BSal or p60BSal deletion-bearing derivatives with *Sal*I and ligation with T4 DNA ligase.

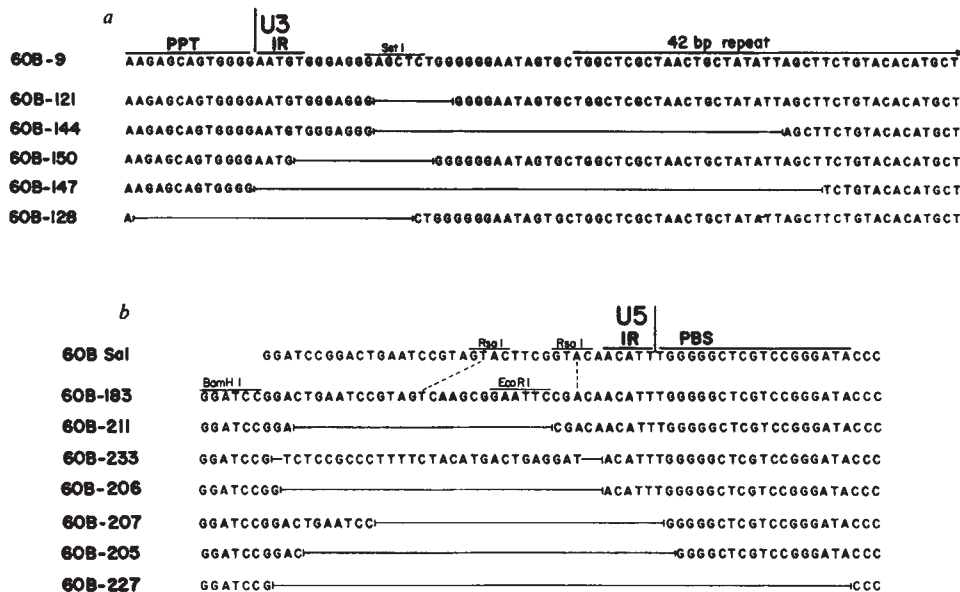
Methods: **a**, For synthesis of deletions at the 5' end of U3, 5 µg of the parental clone p60BSal were digested with 5 U of *Sst*I for 4 min and the reaction was terminated by adding excess EDTA. This incubation gave a partial digestion of the DNA such that the plasmid was cut, on average, at only one of the two *Sst*I sites. The *Sst*I was removed by phenol extraction, the DNA was precipitated with ethanol, the dried DNA was resuspended in 50 mM sodium acetate (pH 4.5), 5 mM ZnSO₄, and *S*₁ nuclease was added to generate blunt-ended DNA; the amount of *S*₁ nuclease was previously titrated so that no digestion of double-stranded DNA would occur. The DNA was rejoined with T4 DNA ligase and used to transform calcium chloride-treated competent *Escherichia coli* cells. Colonies that grew on medium containing 50 µg ml⁻¹ ampicillin were screened by the sodium hydroxide plasmid 'miniprep' procedure⁵⁰. A plasmid lacking the *Sst*I site outside the LTR was identified and designated p60B-9. 25 µg of CsCl-purified p60B-9 were digested to completion with *Sst*I, phenol-extracted and ethanol-precipitated. The dried DNA was resuspended in 250 µl of T4 DNA polymerase buffer⁵¹ containing 100 mM dTTP. 35 U of T4 DNA polymerase were added, the 3'-5' exonuclease reaction was allowed to proceed at 37 °C for 5 min, and the reaction was terminated by incubation at 68 °C for 10 min. Following phenol extraction and ethanol precipitation, the DNA was digested with *S*₁ nuclease to generate blunt ends, treated with T4 DNA ligase and used to transform competent *E. coli* cells. Plasmids bearing short deletions were identified by the miniprep procedure and were prepared on a large scale by CsCl equilibrium centrifugation. To construct deletions at the 3' end of U5, phosphorylated *Hind*III linkers were ligated into a parental plasmid after partial digestion with *Rsa*I. The resulting plasmid was designated p60B-58. The *Hind*III site of p60B-58 was converted to an *Eco*RI site by digestion with *Hind*III, treatment with *S*₁ nuclease and ligation with *Eco*RI linkers, yielding p60B-183. The *Eco*RI site in the pBR322 DNA had been previously removed by cleavage with *Eco*RI and *S*₁ nuclease treatment. Deletions from the unique *Eco*RI site near the end of U5 in p60B-183 were generated in the same way as were the deletions in U3. **b**, A subpopulation of the concatemers contain a complete copy of the SNV genome with two copies of the LTR. Infection with virus recovered from transfected cells results in the synthesis of viral DNA containing sequences identical to that in the concatemer. Thus, a molecule bearing a deletion at the end of U3 results in a virus strain with a deletion at the outside terminus of the 'leftward' LTR and an identical deletion at the inside end of the 'rightward' LTR. Similarly, a deletion at the end of U5 will be located at the outside end of the rightward LTR and the inside end of the leftward LTR.

LTR DNA in the process of integration during spleen necrosis virus (SNV) infection of chicken embryo fibroblasts. SNV is a retrovirus of the avian reticuloendotheliosis species with genus-specific relationships to mammalian type C viruses²⁹⁻³¹. Removal of specific sequences at the extreme termini, including the region containing the 5-bp inverted repeats present in unin-

tegrated DNA, completely abolishes the ability of viral DNA to integrate into the host cell DNA. However, virus strains deficient in the establishment of the proviral state can still produce virus, although in diminished amounts. Thus, integration of retrovirus DNA is not an absolute prerequisite for retrovirus replication.

Fig. 2 DNA sequence analysis of deletions at the ends of the LTR. In *a* a segment of one 42-bp repeat in U3 is depicted. It has been suggested that these repeats have a role in transcription initiation within the LTR. In *b*, comparison of p60BSal and p60B-183 indicates that the 8 bp between two closely spaced *Sst*I sites in p60BSal have been lost and replaced by a substitution containing an *Eco*RI site in p60B-183 as predicted in Fig. 1. The dotted lines between 60BSal and 60B-183 connect the ends of this substitution. The extent of individual deletions is shown by horizontal lines. The 5-bp inverted repeats (IR) at the termini of the LTR, the polypurine tract (PPT), primer-binding site (PBS) and ends of U3 and U5 are shown.

Methods: Individual plasmids generated by exonuclease digestion and *S*₁ treatment, as described in Fig. 1 legend, were subjected to chemical DNA sequence analysis. *a*, For the mutants in U3, a unique *Hae*III fragment containing the deletions was excised from a 5% polyacrylamide gel, and the DNA was eluted in 40 mM Tris, 80 mM boric acid, 2 mM Na₂ EDTA and was collected on a 0.1-ml bed of DEAE-52 in buffer containing 50 mM Tris (pH 7.9), 2 mM EDTA, 0.15 M NaCl. The DNA was removed from the column in buffer containing 1.0 M NaCl, concentrated by ethanol precipitation, and labelled at the 5' ends with [γ -³²P]ATP and T4 polynucleotide kinase using the exchange reaction with ADP³². After cleavage of the fragments at an internal *Hinf*I site, the fragments were repurified by electrophoresis on 8% polyacrylamide gels, recovered from the gels, and the nucleotide sequences were determined by the method of Maxam and Gilbert³². *b*, For analysis of the mutants in U5, the DNA was cleaved with *Bam*HI (there is a *Bam*HI recognition site 28 bp 5' from the *Eco*RI site in p60B-183), labelled at the 5' end, and digested with *Sst*I. The fragments containing the deletions were identified and purified by electrophoresis on 5% polyacrylamide gels, and the DNA sequences were determined.



Synthesis of deletion mutants

The recombinant plasmid p60BSal contains a complete copy of SNV DNA circularly permuted from the single *Sal*I restriction endonuclease site located 0.86 kilobase pairs (kbp) from one end of the viral DNA^{32,33} (Fig. 1a). Consequently, *Sal*I cleavage of the plasmid and head-to-tail joining of viral DNA fragments by T4 DNA ligase results in a concatemeric molecule containing sequences colinear with unintegrated viral DNA (Fig. 1b). Calcium phosphate co-precipitation of DNA treated in this manner followed by transfection of chicken embryo fibroblasts results in the production of infectious SNV.

The strategy used to generate deletion mutants near the ends of the SNV LTRs, including the short 5-bp inverted repeat located at the ends of U3 and U5, is shown schematically in Fig. 1a. Deletions near the 5' end of U3 were synthesized by exonucleolytic excision extending from a naturally occurring *Sst*I restriction site located 14 bp from the end of U3. Similarly, deletions at the 3' terminus of U5 were made extending from an artificially introduced *Eco*RI restriction site located 13 bp from the end of U5. In the parental plasmid, p60BSal, the complete 5-base inverted repeat is present at both ends of the LTR.

As will be shown below, deletion of specific sequences at the termini of the LTR resulted in one of four phenotypes: (1) strains essentially normal for virus production and integration, (2) strains with depressed levels of virus production and no detectable integration, (3) strains totally defective in virus production and integration and (4) a single unusual strain bearing a substitution which is depressed in both virus production and integration.

Nucleotide sequence characterization identified plasmids that lack nucleotides at or near the termini of the LTR. At the 5' end of U3, p60B-121 and -144 contain deletions leaving the terminal 12 bases intact (Fig. 2a). The plasmid p60B-150 contains a deletion leaving only the terminal four nucleotides while p60B-147 bears a deletion of the entire terminus up to the

polypurine tract; the polypurine tract has a putative function in the initiation of plus-strand DNA synthesis³⁴⁻³⁶. Finally, p60B-128 lacks the entire terminus plus 12 nucleotides of the polypurine tract. At the 3' end of U5, p60B-211 and p60B-206 contain deletions up to the terminal 10 and 5 nucleotides, respectively (Fig. 2b). The plasmids p60B-207, -205 and -227 have deletions extending past the U5 terminus and into the primer-binding site; the primer-binding site is required for initiation of minus-strand DNA synthesis by cellular tRNA^{Pro} (refs 21, 37, 38). A second unusual mutation was also identified that contains a 32-base deletion extending to the beginning of the inverted repeat in U5 and a 29-base substitution derived from DNA of unknown origin (p60B-233).

Virus production

DNA cleaved by endonuclease *Sal*I and rejoined by ligation was used to transfect fibroblast cells as described in Fig. 3 legend. Virus production from these transfected cells and from cells subsequently infected with the virus recovered from the supernatant medium from transfected cells was monitored by the presence of acute cytopathic effects and DNA polymerase activity³⁹.

Transfection or infection with some strains bearing deletions of all or part of the 3' terminus of U3 (p60B-150 and -147) or the 5' terminus of U5 (p60B-206) resulted in substantial virus production, although levels were 1,000–10,000 times less than those of the derivatives with deletions ending farther from the termini (p60B-121, -144 and -211) (see Fig. 7). These observations indicate that the termini are not essential for small amounts of virus production. However, deletion of nucleotides past the ends of the LTR into the primer-binding site (p60B-227) or the polypurine tract (p60B-128) led to the loss of detectable virus replication, presumably due to the inability to initiate significant minus- or plus-strand DNA synthesis. The substitution mutation (p60B-233) also produced small amounts of virus and, as described below, viral DNA derived from this strain and

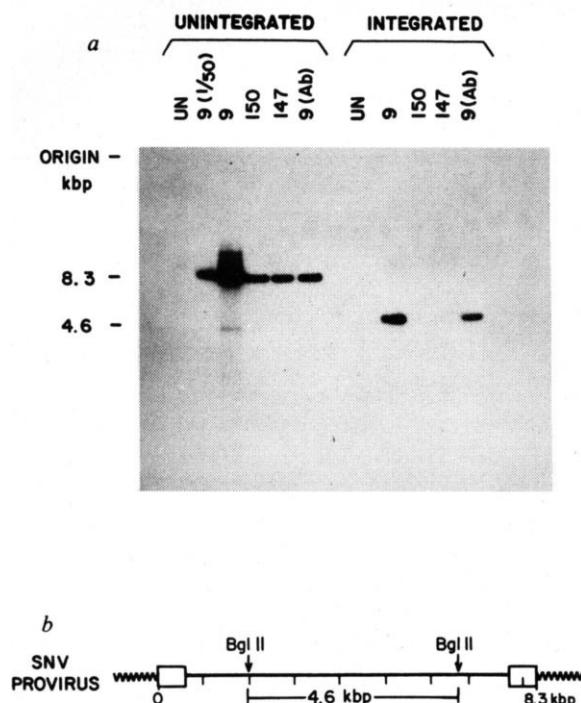


Fig. 3 Unintegrated and integrated viral DNAs synthesized during infection with mutants bearing deletions at the 5' end of U3. Lanes with unintegrated DNA contained DNA from one 100-mm diameter plate or 1/50th of that amount of DNA where indicated [9(1/50)]. Lanes containing integrated DNA each contained 20 μ g of DNA except that the lane from 60B-9-infected cells grown in the absence of antiserum contained 5 μ g DNA. UN is DNA from uninfected cells. **b** Shows the structure of an SNV provirus (integrated DNA). The viral LTRs and surrounding cellular DNA are represented by the wide bars and wavy lines, respectively.

Methods: **a**, The recombinant plasmids p60B-9, -150 and -147 were digested with *SalI*, the enzyme was inactivated with 1% diethylpyrocarbonate, and the DNA was ligated to form concatemers like those depicted in Fig. 1b. The DNA was co-precipitated with calcium phosphate⁵³ in the presence of 10 μ g ml⁻¹ sheared salmon sperm DNA and used to transfect chicken embryo fibroblasts. Virus collected 3 days after transfection was used to infect 2×10^6 chicken embryo cells per 100-mm culture dish in the presence of 10 μ g ml⁻¹ polybrene. After 40 min of incubation, media were added to the plates and incubation was continued for an additional 48 h. In one instance, rabbit antisera to SNV was added to the media to prevent multiple rounds of infection [9(AB)]: antisera were added 6 h after infection. The cells were fractionated into nuclear and cytoplasmic components by a 10-min incubation in the presence of 0.3 ml of an ice-cold solution of 0.1 M NaCl, 10 mM Tris (pH 7.9), 0.5% NP40, 1.5 mM MgCl₂ and a subsequent 10 min centrifugation at 5,000 r.p.m. in a Sorvall SS-34 rotor. The supernatant (cytoplasmic fraction) was extracted with a 1:1 mixture of phenol/chloroform and ethanol-precipitated. The sediment (nuclear fraction) was subjected to the Hirt fractionation procedure⁵⁴. The Hirt supernatant fraction was pooled with the cytoplasmic DNA and designated unintegrated DNA. The DNA of the Hirt pellet fraction extracted from the purified nuclei was designated integrated DNA. Unintegrated DNA was applied to a 1% agarose gel without treatment, and integrated DNA was digested with *BglII* before loading. Following electrophoresis, the DNA was transferred to nitrocellulose by the method of Southern⁵⁵ and hybridized with p60BSal which had been labelled by nick-translation.

from strains lacking only 1 or 2 bases of the primer-binding site (p60B-207, -205) could also be detected by nucleic acid hybridization analysis. The segment of substituting DNA in p60B-233 contains a start codon and is located upstream from the start codon of the *gag* gene. This substitution may have resulted in disruption of normal translation initiation following transfection by this mutant.

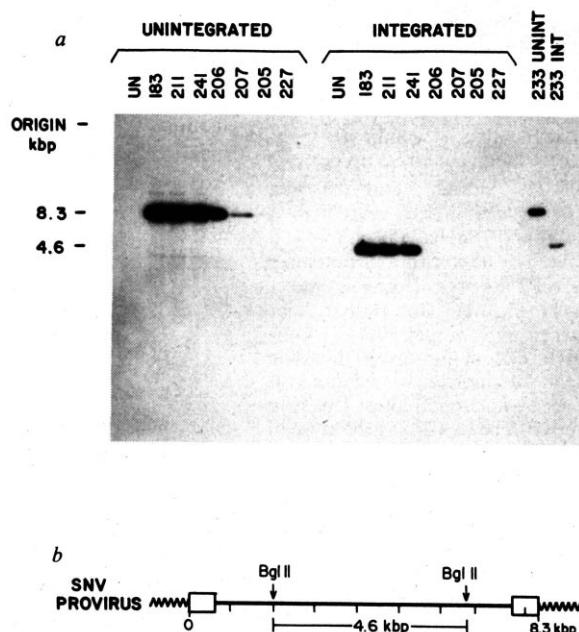


Fig. 4 **a**, Unintegrated and integrated viral DNA synthesized during infection with mutants bearing deletions at the 3' end of U5. **b**, Structure of an SNV provirus as in Fig. 3b. In **a**, methods were the same as described in Fig. 3 legend. Lanes containing unintegrated DNA from cells infected with 60B-206, -207, -205, -227 and -233 contained DNA from one 100-mm diameter plate. Lanes with unintegrated DNA from 60B-183, -211 and -241 contained 1/20th the amount of DNA from plates of the same size. Lanes with integrated DNA from 60B-206, -207, -205, -227 and -233 infected cells contained 20 μ g of DNA and those with DNA from 60B-183, -211 and -241 contained 5 μ g of DNA. p60B-241 has an 18-bp deletion that includes only a subset of the nucleotides deleted from p60B-211.

Integration

The virus strains containing full or partial deletions of the terminal inverted repeats in U3 or U5 were examined for their ability to form unintegrated and integrated viral DNA as described in Fig. 3 legend. Proviral DNA was detected by digesting cellular DNA with *BglII*, which cleaves SNV DNA at two sites, generating a unique 4.6-kbp fragment and two terminal fragments of variable length depending on the location of *BglII* sites in the adjacent cellular DNA. Hence, only the internal fragment is common to all copies of the integrated viral DNA. Unintegrated viral DNA was detected without prior endonuclease digestion.

The results of the analysis were striking. Several viruses containing deletions at the terminus at either end of the LTR were incapable of integrating into the cellular DNA even though unintegrated viral DNA was synthesized (p60B-150, -147 and -206) (Figs 3, 4). However, mutants carrying deletions extending to 7 bases from the inverted repeat in U3 (p60B-121) or 3 bases from the inverted repeat in U5 (p60B-211) were competent for integration. Moreover, the substitution mutant p60B-233, which is altered at the end of U5 leaving only the inverted repeat intact, produced low levels of both unintegrated and integrated viral DNA. This latter observation demonstrates that, at least at the U5 terminus, the bases immediately adjacent to the 5-base inverted repeat are important for integration.

The strains unable to integrate are integration-deficient to the limits of our detection; integration is decreased to $<10^{-5}$ of that present in the integration-competent strain, as demonstrated by hybridization to decreasing amounts of integrated viral DNA (Fig. 5). Additionally, the deletion strains retained the *BglII* sites located within the viral DNA as determined by restriction endonuclease digestion of unintegrated viral DNA (data not shown). Thus, had integration occurred in these

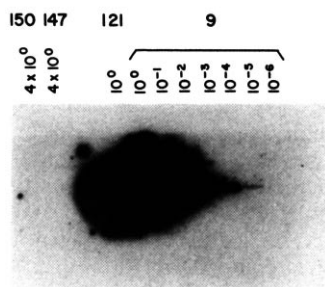


Fig. 5 Quantification of integrated viral DNA. Integrated viral DNA was isolated and digested with *Bgl*II as described in Fig. 3 legend. 20 μ g of DNA from cells infected with the integration-deficient strains 60B-150 and -147 were subjected to electrophoresis in parallel with 5 μ g of DNA from the integration-competent strains 60B-121 and -9 and successive 10-fold dilutions of the latter DNA. The DNA was transferred to nitrocellulose and analysed by hybridization as described in Fig. 3 legend.

mutants, *Bgl*II digestion of the DNA would have resulted in the generation of the 4.6-kbp fragment from the proviral DNA.

We examined unintegrated SNV DNA in the nucleus and cytoplasm of infected cells to determine whether the deletions affected the transport of viral DNA into the nucleus. As shown in Fig. 6, DNA from the integration-deficient strains 60B-150 and -206 can enter the nucleus as efficiently as can DNA of the wild-type strain. In addition, longer exposure of the filter revealed open circular forms of the viral DNA in the nuclear fraction from cells infected with either integration-competent or integration-deficient strains (data not shown). Thus, the deletions affecting the ability of the viral DNA to integrate do not block entry of unintegrated viral DNA into the nucleus or the subsequent circularization of the DNA.

Strains lacking the first or first and second bases of the primer-binding site produced apparently intact unintegrated viral DNA, albeit at greatly diminished levels relative to those strains maintaining the complete primer-binding site (p60B-207 and -205) (Fig. 4). Quantification of the unintegrated viral DNA shown in Fig. 4 reveals that p60B-207 produces only about 1/10th the amount of DNA produced by the integration-deficient strain p60B-206. The removal of the two terminal bases of the primer-binding site (p60B-205) results in a 1,000-fold reduction in unintegrated viral DNA relative to p60B-206. As expected, p60B-207 and -205 are also unable to form a provirus (Fig. 4).

Conclusions

The data presented here clearly show that the termini of the retrovirus LTR are essential for the establishment of the proviral state. More exactly, the termini of the LTR are required for integration, as alteration of this region results in retrovirus strains deficient in integration.

Examination of nucleotide sequences placed in juxtaposition with the LTR termini through the process of deletion reveals that the regions critical for integration are very close to the ends of the LTR (Fig. 7). In U3, the region required for integration lies within 12 bp of the terminus, and the region of U5 needed for integration is within the last 8 bp. However, alteration of these regions alone, at least in U5, does not necessarily block integration; 60B-206 and 60B-233 both contain substituted bases in the four nucleotides adjacent to the terminal inverted repeat, but 60B-233 is integration-competent and 60B-206 is not. Quantification of the unintegrated and integrated viral DNA synthesized after infection by 60B-233 indicates that the efficiency of integration may be only about one-half that observed for other integration-competent virus strains such as 60B-121 and -183.

The mutations present in the integration-deficient virus strains appear to be stably maintained during passage. When integrated

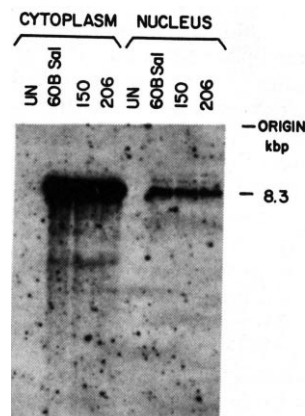


Fig. 6 Distribution of unintegrated viral DNA in the cytoplasm and nucleus of SNV-infected cells. Cells were infected with the parental virus 60B-Sal and the deletion-containing viruses 60B-150 and 60B-206. The latter two viruses are integration-deficient virus strains (Figs 3, 4). After 48 h the cells were divided into the nuclear and cytoplasmic fractions by treatment with buffer containing NP40 as described in Fig. 3 legend. The Hirt supernatant DNA from the purified nuclei and the cytoplasmic DNA were analysed by gel electrophoresis and nucleic acid hybridization as described in Fig. 3 legend. For the lane containing 60B-Sal DNA, 1/50th the amount of DNA used for lanes with 60B-150 and 60B-206 DNA was applied to the gel.

		UNINTEGRATED	INTEGRATED	VIRUS PRODUCTION (IU/ml)
wt		1.0	1.0	10 ⁷
60B-121		0.9	0.9	10 ⁷
60B-150		0.04	<10 ⁻⁵	10 ⁴
60B-147		0.03	<10 ⁻⁵	10 ³
wt		1.0	1.0	10 ⁷
60B-183		0.4	0.4	10 ⁶
60B-233		0.02	0.01	10 ²
60B-206		0.09	<10 ⁻⁵	10 ⁴
60B-207		0.01	<10 ⁻⁵	0*
60B-205		10 ⁻⁵	<10 ⁻⁵	0*

Fig. 7 Nucleotide variation at the U3 and U5 termini of strains bearing deletions. An x over an individual nucleotide indicates that the base deviates from that of the wild-type nucleotide at the same position. Quantification of unintegrated and integrated viral DNA was achieved by scanning the lanes of Figs 3 and 4 with a Beckman DU-8 spectrophotometer (multiple time exposures of the film were used in the analysis). Separate gels were used in the comparison of the DNA from wild type (60B-Sal) and the derivatives 60B-9 and 60B-183 (not shown). For quantification of virus production, cells were transfected as described in Fig. 3 legend, the virus was passaged at least twice, and 10-fold dilutions were used to infect freshly plated cells. Virus production was monitored by the presence of DNA polymerase activity and acute cytopathic effects^{39,40}. *Virus was not detected by DNA polymerase activity or cytopathic effects but unintegrated viral DNA was detected by nucleic acid hybridization (Fig. 4).

viral DNA from cells infected with 60B-150 was subjected to digestion with *Sst*, the *Sst* recognition site was still absent as expected (not shown). Moreover, reversion to the wild-type genotype would have resulted in the generation of integration-competent virus, but the integration-deficient phenotype was maintained through passage.

The experiments described here are the first to indicate that establishment of the provirus is not required for retrovirus replication. Unintegrated viral DNA contains all sequences

needed for expression of the retrovirus genes, and this form can serve as a suitable template for transcription. Virus production from the integration-deficient virus strains is lower than that observed for integration-competent strains (Fig. 7). These results indicate that unintegrated viral DNA is an inferior transcriptional template. Further, if the DNA conformation is important for transcription, it may be that only the small minority of unintegrated viral DNA in the circular form may serve as an efficient template. It is unlikely that virus production is derived from rare copies of integrated DNA from the integration-deficient viruses. Integration is decreased by at least 10^{-5} of that found in wild-type virus but virus production is diminished by only 10^{-3} . Therefore, if rare integrated viral DNA served as the sole transcriptional template, the level of transcription would have to be increased by $\sim 10^2$. Furthermore, transfer of single cells infected by integration-deficient virus results in a new population of infected cells (data not shown). Thus, the unintegrated viral DNA, not a rare integrated copy, is used as the transcriptional template during infection by integration-deficient virus.

One consequence of SNV infection of chicken embryo fibroblasts is the generation of acute cytopathic effects visible by gross microscopic observation^{8,39-41}. As the integration-deficient viruses produced some cytopathic effects, integration cannot be the cause of the cytopathic effects. It is more likely that the cytopathic effects are a manifestation of the synthesis of a specific viral gene product or perhaps the entrance of virus into the cell.

During initiation of minus-strand DNA synthesis, the primer (cellular tRNA^{Pro}) is hydrogen bonded with the primer-binding site adjacent to U5 (refs 21, 35). Our results indicate that complete deletion of the primer-binding site leads to the loss of detectable unintegrated or integrated viral DNA. However, the removal of one or two bases at the 5' end of the primer-binding site still allows the synthesis of unintegrated viral DNA although at greatly diminished levels. This indicates that some initiation of minus-strand synthesis can occur even though the 3' end of the primer (3'-AC) is not complementary to the template. Alternatively, the primer could be a tRNA derivative lacking these bases.

Complete removal of the polypurine tract adjacent to U3 also results in the loss of detectable viral DNA, in agreement with previously observed results³⁸. Like the initiation of minus-strand synthesis, the initiation of plus-strand synthesis is precise³⁷. Because a primer for plus-strand DNA synthesis has not been

identified, it has been speculated that specific endonucleolytic digestion of the viral RNA by RNase H could provide such a primer^{37,42}. Deletion of the entire inverted repeat of U3 does not block viral DNA synthesis (strain 60B-147), indicating that endonucleolytic cleavage probably depends only on the nucleotides that comprise the adjacent polypurine tract.

There are structural similarities between the genome of the DNA-containing hepatitis B viruses and that of retrovirus DNA⁴³. Moreover, replication of duck hepatitis B virus involves an RNA intermediate and DNA synthesis catalysed, at least in part, by an RNA-dependent DNA polymerase activity⁴⁴. These features indicate that the retroviruses and hepatitis B viruses may be evolutionarily related⁴⁵. The integration-deficient retroviruses described here share an additional feature with hepatitis B: integration of the viral DNA into the genome of the host cell is not a mandatory step in the viral replicative cycle.

Available data indicate that there is an evolutionary relationship between transposable elements and retroviruses (the provirus hypothesis)^{21,23,25,27,28,45,46}. Studies with two prokaryotic transposable elements indicate that *cis*-acting sequences required for transposition are located at the termini. In Tn10, the outermost 70 bp at each end of the element appear to have a *cis*-acting function in transposition⁴⁷. Recent work with derivatives of Tn5-bearing deletions approaching the termini show that, at least at one end of the transposon (IS50L), only the 16-18 bp at the end of the element are required for transposition⁴⁸. Tn5 contains a 9-bp terminal inverted repeat⁴⁹.

Integration of retrovirus DNA poses different structural problems from transposition of prokaryotic transposable elements: integration of viral DNA occurs from a molecule with either two free ends or a circular derivative, while transposition results in relocation and concomitant duplication of a segment of DNA already bounded by adjacent host DNA. It will be interesting to see whether the parallels in nucleotide sequence organization at the termini of retrovirus DNA and DNA of transposable elements reflect similarities in the enzymatic mechanisms of integration and transposition.

We thank D. Dillman, K. Eggleston, S. Hellenbrand and A. Joy for technical assistance and J. Embretson, M. Emerman, C. Miller, B. Sugden, G. Tarpley and K. Wilhelmson for comments on the manuscript. This work was supported by NCI grants CA-22443 and CA-07175. A.T.P. is a postdoctoral fellow of the Damon Runyon-Walter Winchell Cancer Fund. H.M.T. is an American Cancer Society Research Professor.

Received 9 June; accepted 24 August 1983.

1. Fritsch, E. & Temin, H. M. *J. Virol.* **21**, 119-130 (1977).
2. Shank, P. R. & Varmus, H. E. *J. Virol.* **25**, 104-114 (1978).
3. Guntaka, R. V. *et al. J. molec. Biol.* **106**, 337-357 (1976).
4. Khoury, A. T. & Hanafusa, H. *J. Virol.* **18**, 383-400 (1976).
5. Varmus, H. E., Heasley, S., Linn, J. & Wheeler, K. J. *J. Virol.* **18**, 574-585 (1976).
6. Battula, N. & Temin, H. M. *Cell* **13**, 387-398 (1978).
7. Hsu, T. W., Sabran, J. L., Mark, G. E., Guntaka, R. V. & Taylor, J. M. *J. Virol.* **28**, 810-818 (1978).
8. Shank, P. R. *et al. Cell* **15**, 1383-1395 (1978).
9. Varmus, H. E. *et al. J. molec. Biol.* **120**, 55-82 (1978).
10. Gilboa, E., Mitra, S. W., Goff, S. & Baltimore, D. *Cell* **18**, 93-100 (1979).
11. Kung, H.-J., Shank, P. R., Bishop, J. M. & Varmus, H. E. *Virology* **103**, 425-433 (1980).
12. Shimotohno, K. & Temin, H. M. *J. Virol.* **41**, 163-171 (1982).
13. Even, J. *et al. J. Virol.* **45**, 1004-1016 (1983).
14. Hughes, S. E. *et al. Cell* **15**, 1397-1410 (1978).
15. Keshet, E., O'Rear, J. J. & Temin, H. M. *Cell* **16**, 51-61 (1979).
16. Sabran, J. L. *et al. J. Virol.* **29**, 170-178 (1979).
17. Scott, M. L., McKereghan, K., Kaplan, H. S. & Fry, K. E. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4213-4217 (1981).
18. Ju, G. & Skalka, A. M. *Cell* **22**, 379-386 (1980).
19. Yamamoto, T., Jay, G. & Pastan, I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 176-180 (1980).
20. Sutcliffe, J. G., Shinnik, T. M., Verma, I. M. & Lerner, R. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3302-3306 (1980).
21. Shimotohno, K., Mizutani, S. & Temin, H. M. *Nature* **285**, 550-554 (1980).
22. Donehower, L. A., Huang, A. L. & Hagen, G. L. *J. Virol.* **37**, 226-238 (1981).
23. Dhar, R., McClements, W. L., Enquist, L. W. & Vande Woude, G. F. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3937-3941 (1980).
24. Hagen, G. L. & Donehower, L. A. *UCN-UCLA Symp. molec. cell. Biol.* **18**, 255 (1980).
25. Shimotohno, K. & Temin, H. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7357-7361 (1980).
26. Van Beveren, C., Goddard, J. G., Berns, A. & Verma, I. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3307-3311 (1980).
27. Majors, J. E. & Varmus, H. E. *Nature* **289**, 253-258 (1981).

28. Hughes, S. H., Mutschler, A., Bishop, J. M. & Varmus, H. E. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4299-4303 (1981).
29. Hunter, E., Bhowan, A. S. & Bennet, J. C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2708-2712 (1978).
30. Barbacid, M., Hunter, E. & Aaronson, S. A. *J. Virol.* **30**, 508-514 (1979).
31. Bauer, G. & Temin, H. M. *J. Virol.* **34**, 168-177 (1980).
32. O'Rear, J. J., Mizutani, S., Hoffman, G., Fiant, M. & Temin, H. M. *Cell* **20**, 423-430 (1980).
33. Watanabe, S. & Temin, H. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5986-5990 (1982).
34. Czernilofsky, A. P. *et al. Nucleic Acids Res.* **8**, 2967-2984 (1980).
35. Mitra, S. W., Chow, M., Champoux, J. & Baltimore, D. *J. biol. Chem.* **257**, 5983-5986 (1982).
36. Sorge, J. & Hughes, S. H. *J. Virol.* **43**, 482-488 (1982).
37. Peters, G. *et al. J. Virol.* **21**, 1031-1041 (1977).
38. Peters, G. & Dahlberg, J. E. *J. Virol.* **31**, 398-407 (1979).
39. Temin, H. M. & Kassner, V. K. *J. Virol.* **13**, 291-297 (1974).
40. Temin, H. M. & Kassner, V. K. *J. gen. Virol.* **27**, 267-274 (1975).
41. Keshet, E. & Temin, H. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3372-3376 (1978).
42. Omer, C. A. & Faras, A. J. *Cell* **30**, 797-805 (1982).
43. Summers, J. & Mason, W. S. *Hepatology* **2**, 165-665 (1982).
44. Summers, J. & Mason, W. S. *Cell* **29**, 403-415 (1982).
45. Temin, H. M. *Cell* **21**, 599-600 (1980).
46. Temin, H. M. *Perspectives Biol. Med.* **14**, 11-26 (1970).
47. Foster, T. J., Davis, M. A., Roberts, D. E., Takeshita, K. & Kleckner, N. *Cell* **23**, 201-213 (1981).
48. Johnson, R. C. & Reznikoff, W. S. *Nature* **304**, 280-282 (1982).
49. Berg, D. E., Johnsrud, L., McDivitt, L., Ramabhadran, R. & Hirschel, B. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2632-2635 (1982).
50. Birnboim, H. C. & Doly, J. *Nucleic Acids Res.* **7**, 1513-1523 (1979).
51. O'Farrell, P. *Bethesda Res. Labs Focus* **3**, 1 (1981).
52. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-520 (1980).
53. Graham, F. L. & Van der Eb, A. J. *Virology* **52**, 456-467 (1973).
54. Hirt, B. *J. molec. Biol.* **26**, 365-369 (1967).
55. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).

The end may be hastened by magnetic monopoles

Michael S. Turner

Astronomy and Astrophysics Center, Enrico Fermi Institute,
The University of Chicago, Illinois 60637, USA

In an open Universe, or even in just a long-lived Universe (as predicted in inflationary Universe models¹⁻³), the ultimate demise of matter is a certainty. Stars, of course, have only a fleeting existence as they must eventually exhaust their nuclear fuels (in $\sim 10^{10}$ yr for a $1 M_{\odot}$ star). In 10^{14} yr or so new star formation will have ceased, and only neutron stars, white dwarfs, planetary-sized objects (such as Jupiter and Earth) and black holes will remain⁴. In the simplest grand unified theories (GUTs) the proton lifetime is but $0(10^{30})$ yr, so that after this time neutron stars, white dwarfs, and planets will cease to exist. (Due to gravitational effects alone the proton lifetime is expected to be $0(10^{46})$ yr or less.) Although they could outlive the other three, black holes will also eventually disappear, evaporating through the Hawking process⁵ in a time $\approx 10^{64}(M/M_{\odot})^3$ yr. I point out here that the existence of just a tiny flux of superheavy magnetic monopoles, say $F \approx F_{-21} \times 10^{-21} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$, will, due to monopole-catalysed nucleon decay^{6,7}, hasten the demise of matter. Due to the monopoles that collect inside them, the lifetime of a neutron star will be $\approx 10^{11} F_{-21}^{-1/2}$ yr, of a white dwarf $\approx 10^{14} F_{-21}^{-1/2}$ yr, and of an Earth-like planet $\approx 10^{18} F_{-21}^{-1/2}$ yr.

GUTs not only predict the instability of the proton, but also usually predict the existence of stable, superheavy magnetic monopoles. In 1974 't Hooft⁸ and Polyakov⁹ showed that when a semi-simple group G breaks down to a smaller group G' which contains a $U(1)$ factor there exist stable, topological solitons in the theory, whose long range fields are those of a Dirac monopole¹⁰, and whose mass (the energy associated with the configuration) m is $\approx M_G/\alpha$ where α is the unified coupling constant and M_G is the energy scale of symmetry breakdown. For $SU(5)$, symmetry breakdown to $SU(3) \times SU(2) \times U(1)$ occurs at an energy scale of $0(10^{14})$ GeV, and the superheavy monopole mass is predicted to be $0(10^{16})$ GeV $\approx 10^{-8}$ g. (For definiteness I shall use the $SU(5)$ monopole for all further discussions.) At large distances ($r \gg M_G^{-1} \approx 10^{-28}$ cm) these GUT monopoles closely resemble a Dirac monopole. However, at small distances ($r \lesssim M_G^{-1}$) the differences between a Dirac monopole and a 't Hooft-Polyakov monopoles are profound. At the core of the monopole ($r \approx M_G^{-1}$) the vacuum expectation value of the scalar field which breaks G down to G' vanishes, and the full symmetry of the theory is manifest. If the theory does not conserve baryon number, then in the core, where the gauge bosons which mediate such processes are massless, baryon number nonconserving processes proceed at the same rate as all other gauge interactions.

With such a structure, one would expect monopoles to catalyse nucleon decay¹¹ (for example, $M + p \rightarrow M + \pi^0 + e^+$), with a cross-section of the order of the geometric cross-section of the core, $M_G^{-2} \approx 10^{-56} \text{ cm}^2$ —an utterly negligible cross-section. Quite remarkably, however, Callan⁶ and Rubakov⁷ have shown that the cross-section for monopole-catalysed nucleon decay should be orders of magnitude larger, comparable in size to a 'typical hadronic cross-section' (say 10^{-27} cm^2). Their result relies on the singular nature of the potential between a Dirac monopole and a fermion, $V(r) \propto r^{-2}$. For s-wave scattering the wave function of the fermion is literally 'sucked into' the core, resulting in a cross-section which saturates the unitarity limit $\sigma \propto E^{-2}$. The large catalysis cross-section can have profound effect on the ultimate fate of matter in the Universe.

Clearly there is no contemporary site (terrestrial, astrophysical or otherwise) where particles of mass 10^{16} GeV can be produced. Any superheavy monopoles present today must be relics from the very earliest moments of the Universe ($t \approx 10^{-35}$ s). Thus far, all attempts to calculate their relic abundance have been unsatisfactory in some regard. In the standard big-bang cosmology, their predicted relic abundance is unacceptably large^{12,13} (resulting in a model Universe which is only 30,000 yr old when $T \approx 3$ K). In the inflationary Universe scenarios¹⁻³, they are only produced in rare, very energetic particle collisions, and their predicted relic abundance is $\propto \exp(-2m/T_{RH})$ (T_{RH} = temperature to which the Universe is reheated after inflation)¹⁴. The quantity $2m/T_{RH}$ is model-dependent, but typically $0(\text{few } \alpha^{-1}) \approx \text{few} \times 10^2$. For $SU(5)$ $2m/T_{RH} \approx 400$, which implies a discouragingly tiny monopole abundance— < 1 in the observable Universe.

In the absence of a firm prediction for the monopole abundance let us suppose that the present flux of monopoles is

$$F = F_{-21} 10^{-21} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} \quad (1)$$

The reason for this parameterization will become clear later. A flux of $10^{-21} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$ is some 11 orders of magnitude smaller than the 'flux' inferred from Cabrera's candidate event¹⁵ and some six orders of magnitude smaller than the Parker bound¹⁶, and corresponds to about one monopole passing through a large city (size ≈ 30 km) per year. The velocity of monopoles relative to astrophysical objects (galaxies, stars, planets and so on) should be of the order of $10^{-3} c$, because (to within an order of magnitude) this is the virial velocity within a galaxy or cluster of galaxies, the typical peculiar velocity of a galaxy, and the velocity to which a monopole is accelerated by the galactic field¹⁷. Using $10^{-3} c$ as the typical monopole velocity, it follows that an average monopole flux F_{-21} corresponds to an average monopole-to-nucleon ratio of about $2 \times 10^{-21} F_{-21}$.

Nucleons exposed to a monopole flux F_{-21} will have a lifetime of¹⁸

$$\tau \approx (4\pi\sigma_{AB}F)^{-1} \approx 2.5 \times 10^{36} F_{-21}^{-1} \text{ yr} \quad (2)$$

where here (and throughout) the cross-section for catalysis times monopole-nucleon relative velocity is taken to be constant⁷ and equal to $\sigma_{AB} \approx (10^{-27} \text{ cm}^2) \times (3 \times 10^{10} \text{ cm s}^{-1})$. For $F_{-21} \leq 10^6$ this lifetime is likely to be longer than the lifetimes due to free decay.

However, monopoles which impinge on objects such as stars, planets, and so on can lose energy and be stopped in these objects. Thus monopoles should collect inside stars and planets, thereby greatly enhancing the rate of catalysed nucleon decay. The number of monopoles (N_M) captured by an object of radius $R = R_6 \times 10^6$ cm and mass M_1 solar masses in a time interval $\tau = \tau_{10} \times 10^{10}$ yr is

$$N_M = 4\pi R^2 (\pi - sr) (1 + 2GM/RV_M^2) F \tau \epsilon \\ \approx 1.3 \times 10^{10} R_6^2 F_{-21} \tau_{10} \epsilon (1 + 3 \times 10^5 M_1 R_6^{-1}) \quad (3)$$

where $V_M \approx 10^{-3} c$ the velocity of the monopole far from the object and ϵ is the fraction of monopoles which strike the object and are eventually captured. The factor $(1 + 2GM/RV_M^2)$ is just the ratio of the gravitational capture cross-section to the geometric cross-section.

Monopoles of mass 10^{16} GeV with initial velocities of $10^{-3} c$ will lose sufficient energy in Jupiter-sized objects or larger (for example, Jupiter, white dwarfs and neutron stars) to be stopped¹⁹⁻²⁴, and for these objects $\epsilon \approx 0(1)$. For smaller objects (such as the Earth) only those monopoles initially moving more slowly than some maximum speed V_{MAX} will be stopped (for the Earth²⁰, $V_{MAX} \approx 3 \times 10^{-5} c (m/10^{16} \text{ GeV})^{1/2}$). For these objects ϵ is the fraction of the monopole flux due to monopoles moving more slowly than V_{MAX} , and ϵ should be the order of $(V_{MAX}/10^{-3} c)^4$, which for the Earth is about 10^{-6} (ref. 20).

The captured monopoles inside a body will catalyse nucleon decays, decreasing the number of nucleons N_n at a rate

$$\begin{aligned} dN_n/dt &= -N_n n_n (\sigma_{AB} v) \\ &= -1.1 \times 10^{32} \text{ s}^{-1} M_1 R_6^{-1} F_{-21}^{-1/2} \tau_{10} \epsilon \\ &\quad \times (1 + 3 \times 10^5 M_1 R_6^{-1}) \end{aligned} \quad (4)$$

where n_n is the average number density of nucleons ($\approx 2.9 \times 10^{38} \text{ cm}^{-3} M_1 R_6^{-3}$). The energy release ($\approx 1.5 \times 10^{-3} \text{ erg per nucleon}$) results in a luminosity L and a characteristic temperature T given by

$$\begin{aligned} L &\approx 1.6 \times 10^{29} \text{ erg s}^{-1} M_1 R_6^{-1} F_{-21}^{-1/2} \tau_{10} \epsilon \\ &\quad \times (1 + 3 \times 10^5 M_1 R_6^{-1}) \end{aligned} \quad (5a)$$

$$\begin{aligned} T &\approx (L/4\pi R^2 \sigma)^{1/4} \\ &\approx 1.2 \times 10^5 \text{ K } M_1^{1/4} R_6^{-3/4} F_{-21}^{1/4} \tau_{10}^{1/4} \epsilon^{1/4} \\ &\quad \times (1 + 3 \times 10^5 M_1 R_6^{-1})^{1/4} \end{aligned} \quad (5b)$$

where T was calculated by assuming that the energy is thermalized, and radiated entirely in photons. The predicted luminosity for neutron stars has been compared with observations of discrete (soft) X-ray sources and with measurements of the diffuse (soft) X-ray background to constrain the monopole flux to be $21-24$: $F_{-21} \leq 0(1)$ (which is the reason for the parametrization of the monopole flux chosen here).

Using equation (4) we can estimate the time required for monopoles to eat their way through an object: $\tau_D \approx -N_n/(dN_n/dt)$,

$$\tau_D \approx 6.0 \times 10^{13} \text{ yr } R_6^{1/2} F_{-21}^{-1/2} \epsilon^{-1/2} (1 + 3 \times 10^5 M_1 R_6^{-1})^{-1/2} \quad (6)$$

From equations (5a) and (5b) it follows that the luminosity and temperature of the object during the final stages of its demise ($t \approx \tau_D$) are:

$$\begin{aligned} L_D &\approx 9.5 \times 10^{32} \text{ erg s}^{-1} M_1 R_6^{-1/2} F_{-21}^{1/2} \epsilon^{1/2} \\ &\quad \times (1 + 3 \times 10^5 M_1 R_6^{-1})^{1/2} \end{aligned} \quad (7a)$$

$$\begin{aligned} T_D &\approx 1.1 \times 10^6 \text{ K } M_1^{1/4} R_6^{-5/8} F_{-21}^{1/8} \epsilon^{1/8} \\ &\quad \times (1 + 3 \times 10^5 M_1 R_6^{-1})^{1/8} \end{aligned} \quad (7b)$$

Now consider the fate of the various repositories of nucleons: neutron stars, white dwarfs and planets. A neutron star ($M_1 \approx 1$, $R_6 \approx 1$) will evaporate in a time $\tau_{NS} \approx 1.1 \times 10^{11} F_{-21}^{-1/2} \text{ yr}$, and have a final luminosity and temperature: $L_{NS} \approx 5.3 \times 10^{35} \text{ erg s}^{-1} F_{-21}^{1/2}$ and $T_{NS} \approx 5.1 \times 10^6 \text{ K } F_{-21}^{1/8}$ —which could be easily detected by an X-ray instrument similar to HEAO 2 (Einstein) (should one be operating). Since stable neutron star configurations do not exist for masses $\leq 0.1 M_\odot$, once the mass has decreased to $\approx 0.1 M_\odot$ the neutron star may, in its death throes, explode, producing a display similar to the supernova fireworks which are believed to be associated with the birth of a neutron star²⁵.

A white dwarf ($M_1 \approx 1$, $R_6 \approx 720$) will evaporate in a time $\tau_{WD} \approx 7.9 \times 10^{13} F_{-21}^{-1/2} \text{ yr}$, and have a final luminosity and temperature: $L_{WD} \approx 7.2 \times 10^{32} \text{ erg s}^{-1} F_{-21}^{1/2}$ and $T_{WD} \approx 3.65 \times 10^4 \text{ K } F_{-21}^{1/8}$. This is the fate in store for our Sun. Note that the final luminosity is $\sim 20\%$ of the present solar luminosity; however, the temperature is about a factor of 10 higher.

For a Jupiter-size planet ($M_1 \approx 10^{-3}$, $R_6 \approx 7,130$) we have $\tau_J \approx 5.1 \times 10^{15} F_{-21}^{-1/2} \text{ yr}$, $L_J \approx 1.1 \times 10^{28} \text{ erg s}^{-1} F_{-21}^{1/2}$, and a final

surface temperature of about $730 \text{ K } F_{-21}^{1/8}$ (present surface temperature $\approx 130 \text{ K}$).

Finally, consider the fate of Earth. Using $\epsilon \approx 10^{-6}$, we obtain $\tau_E \approx 1.5 \times 10^{18} F_{-21}^{-1/2} \text{ yr}$, $L_E \approx 3.8 \times 10^{28} \text{ erg s}^{-1} F_{-21}^{1/2}$ and $T_E \approx 140 \text{ K } F_{-21}^{1/8}$. For comparison, the present total energy flux from inside the Earth (due primarily to radioactive K, Th and U decays) is $\sim 3 \times 10^{20} \text{ erg s}^{-1}$. We can take heart in the fact that our Earth will outlive the Sun and most of the planets in the Solar System (Pluto, Mercury and Mars should outlast the Earth, while our race with Venus is very close; of course, the Sun's red giant phase will probably destroy Mercury and Venus).

In making these predictions I implicitly assumed that the monopole flux remains constant. If monopoles are gravitationally bound to galaxies, clusters or superclusters then this is a reasonable assumption. Because of the galactic magnetic field it is doubtful that monopoles could remain bound to the galaxy¹⁷, however, they would probably be bound and could remain bound to clusters and superclusters. If, on the other hand, monopoles are just uniformly distributed throughout space, then due to the expansion of the Universe, their number density decreases as R^{-3} and their flux as R^{-4} (R = cosmic scale factor). Such a rapid decrease implies that objects will have, by the present epoch, captured essentially all of the monopoles they will ever capture. In this case the τ_{10} s in equations (3)–(5) should be replaced by a number of 0(1). The epoch of demise then becomes

$$\tau_D \approx 3.6 \times 10^{17} \text{ yr } R_6 F_{-21}^{-1} \epsilon^{-1} (1 + 3 \times 10^5 M_1 R_6^{-1})^{-1} \quad (8)$$

which results in a bit of a reprieve: $\tau_{NS} \approx 1.2 \times 10^{12} F_{-21}^{-1} \text{ yr}$, $\tau_{WD} \approx 6.2 \times 10^{17} F_{-21}^{-1} \text{ yr}$, $\tau_J \approx 2.6 \times 10^{21} F_{-21}^{-1} \text{ yr}$, and $\tau_E \approx 2.3 \times 10^{26} F_{-21}^{-1} \text{ yr}$.

Is there any escaping this fate? Given the large number of monopoles already accreted, it is too late to worry about constructing monopole shields or deflecting incoming monopoles, unless those monopoles already captured can be extracted and dispersed. The objects with the longest lifetimes are small planets (for which ϵ is small) which are drifting between superclusters where the monopole flux is being diluted as R^{-4} . Of course, there is always the possibility that the monopole flux is exceedingly small or that the catalysis process proceeds at a rate much, much less than is predicted.

This work was supported in part by the DOE through contract AC02-80ER-10773 A003 (at Chicago).

Received 13 June; accepted 30 August 1983.

- Guth, A. *Phys. Rev. D* **23**, 347 (1981).
- Linde, A. *Phys. Lett.* **108B**, 389 (1982).
- Albrecht, A. & Steinhardt, P. *Phys. Rev. Lett.* **48**, 1220 (1982).
- Dyson, F. *Rev. mod. Phys.* **51**, 447 (1979).
- Hawking, S. W. *Nature* **248**, 30 (1974).
- Callan, C. G. *Phys. Rev. D* **25**, 2141 (1982); **D26**, 2058 (1982).
- Rubakov, V. A. *Pis'ma Zh. eksp. teor. Fiz.* **33**, 658 (1981) [*J. exp. theor. Phys. Lett.* **33**, 644 (1981)]. *Nucl. Phys. B* **203**, 311 (1982).
- 't Hooft, G. *Nucl. Phys. B* **79**, 276 (1974).
- Polyakov, A. M. *Pis'ma Zh. eksp. teor. Fiz.* **20**, 430 (1974) [*J. exp. theor. Phys. Lett.* **20**, 194 (1974)].
- Dirac, P. A. M. *Proc. phys. Soc. Lond.* **A133**, 60 (1931).
- Wilczek, F. *Phys. Rev. Lett.* **48**, 1146 (1982).
- Preskill, J. *Phys. Rev. Lett.* **43**, 1365 (1979).
- Zel'dovich, Ya. B. & Khlopov, M. Yu. *Phys. Lett.* **79B**, 239 (1978).
- Turner, M. S. *Phys. Lett.* **115B**, 95 (1982).
- Cabrera, B. *Phys. Rev. Lett.* **48**, 1378 (1982).
- Parker, E. N. *Astrophys. J.* **160**, 383 (1970).
- Turner, M. S., Parker, E. N. & Bogdan, T. *Phys. Rev. D* **26**, 1296 (1982).
- Ellis, J., Nanopoulos, D. V. & Olive, K. A. *Phys. Lett.* **116B**, 127 (1982).
- Ahlen, S. P. in *Magnetic Monopoles* (eds Carrigan, R. A. & Trower, W. P.) (Plenum, New York, 1983).
- Turner, M. S. *Nature* **302**, 804 (1983).
- Kolb, E. W., Colgate, S. A. & Harvey, J. *Phys. Rev. Lett.* **49**, 1373 (1982).
- Dimopoulos, S., Preskill, J. & Wilczek, F. *Phys. Lett.* **119B**, 320 (1982).
- Bais, F. A., Ellis, J., Nanopoulos, D. V. & Olive, K. A. *Nucl. Phys. B* **219**, 189–219 (1983).
- Freese, K. T., Turner, M. S. & Schramm, D. N. *Monopole Catalysis of Nucleon Decay in Old Pulsars*, Preprint (University of Chicago, 1983).
- Page, D. N. *Phys. Lett.* **91A**, 201 (1982).

A search for pulsed IR and near-IR emission from PSR1937+214

J. Middleditch*, D. Cudaback†‡, B. Oliver‡, C. Pennypacker§, M. J. Lebofsky||, G. H. Rieke||, J. T. McGraw||, D. Dearborn||, W. Wisniewski¶ & R. Chini#

* Earth and Space Sciences Division, Los Alamos National Laboratory, University of California

† Radio Astronomy Laboratory, ‡ Department of Astronomy,

§ Space Sciences Laboratory and Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720, USA

|| Steward Observatory, ¶ Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA

Max-Planck-Institut für Radioastronomie, 5300 Bonn 1, Auf dem Hügel 69, FRG

Backer *et al.*^{1,2} discovered PSR1937+214 at radio wavelengths. Djorgovski³ identified a candidate optical counterpart with a red-sensitive CCD on a 1-m telescope, although he found nothing on the paper prints of the Palomar Sky Survey. Manchester *et al.*⁴ at the Anglo-Australian Telescope reported optical pulses from a field including Djorgovski's red candidate, although they have not confirmed this. Their 3.5 arc s field barely included the pulsar position. Two of us (M.J.L. and G.H.R.) have found a bright IR flux of 12.3 mag at 2.2 μ m, an image at the plate limit on the red-sensitive glass print of the Sky Survey, and CO absorption bands, all for the red candidate⁵. We report here that it is probably an ordinary red giant. We do not find statistically significant pulsations at the pulsar frequency or any harmonic up to the fourth in 6 to 8 arc s fields including both the red candidate and the radio pulsar positions.

Table 1 gives all positions known to us. We now conclude that the position difference between the red candidate and the pulsar is real with 2.1 arc s nominal separation, although we did not believe that at the time of the observations. For the red candidate the estimated red magnitude on the Sky Survey glass print is 20 ± 0.5 . Conventional time averaged photometry with the MMT (Multiple Mirror Telescope) and the 2.3-m Steward telescope give $J = 14.2$, $H = 13.0$, $K = 12.3$. Narrowband measurements at 2.2 and 2.33 μ m give a CO index⁶ of $+0.19$ discussed below. Independent evidence of the stellar nature of the red candidate agreeing with our results will be reported by Djorgovski and Spinrad (in preparation).

We searched for pulsations with sufficiently large diaphragms to include all positions in Table 1. Our sensitivity to pulsed light was greater than most measurements of time averaged light, so we could have detected pulsations in directions where an image is not visible. The telescopes, diaphragm apertures, wavelength bands and detectors used for pulsed light searches are: Catalina 1.5-m, 6.5 arc s, unfiltered, Varian III-V; Steward 2.4-m, 7.8 arc s, 1.4–2.3 μ m, He-cooled InSb; Mt Hopkins MMT, 6 arc s, unfiltered, RCA 31034 GaAs. For all observations a well understood system⁷ recorded the number of photomultiplier pulses or voltage-to-frequency converter pulses (following IR detectors) in every 100 μ s interval. The frequency response of the detector system used at the Steward 2.3-m was improved by reducing its preamplifier feedback until its 3-db rolloff was near 1 kHz. The 2.3-m and MMT observations also used a multiscaler display synchronized by a stable frequency synthesizer at the expected apparent pulsar rate.

No significant peaks were seen in the multiscaler display as the data were collected. The recorded time series data were Fourier transformed in groups of 2^{24} or 2^{25} points and the regions of the resulting power spectra bracketing the first four harmonics of the pulsar's apparent frequency were examined to determine the significance of any peaks. Table 2 gives relative

Table 1 Positions of PSR1937+214 and red candidate (1950)

Observer	RA	Dec
PSR1937+214		
Blacker <i>et al.</i> ² VLA	19 h 37 min 28.72 s	21° 28' 1.3"
Becker and Helfand ¹¹ , VLA	19 37 28.74	21 28 1.8
Mean radio pulsar	19 37 28.73	21 28 1.5
Red candidate		
This paper, Sky Survey	19 h 37 min	21° 28' 2.2"
This paper, Steward 2.3-m	19 37 28.85	21 28 3.2
Manchester <i>et al.</i> ⁴ , AAT 4-m	19 37 28.78	21 28 3.3
Mean red candidate	19 37 28.84	21 28 2.9

Table 2 Results of searches for pulsation

Date 1982	15 November	3 December	4 December	8 December
Telescope	1.5-m	2.3-m	2.3-m	MMT
Wavelength	Red	IR	IR	Red
Integration time	50 min	42 min	30 min	52 min
Relative power (limiting magnitude)				
1st Harmonic	0.2(22)	0.2(16.9)	0.3(16.7)	0.3(24.2)
2nd Harmonic	0.4(21.7)	3.7(15.6)	1.8(15.9)	2.7(23.3)
3rd Harmonic	0.8(21.3)	0.2(16.9)	0.7(16.2)	1.9(23.4)
4th Harmonic	0.2(22)	0.3(16.7)	1.6(16)	0.3(24.2)

spectral power levels and limiting stellar magnitudes for 100% modulation concisely at the expected pulse frequencies. The resolution of the power spectra is 0.3 or 0.6 mHz; higher power levels are in other bins but are not considered as candidates for pulsation. The power levels are relative to noise power at nearby frequencies. The red magnitude estimates come from scaling similar measurements made on other telescopes and are uncertain by 0.5 mag. The IR data were calibrated from the measured frequency response of the detector system and by using the chopping secondary of the 2.3-m to modulate the signal from a standard star at 30 Hz. This signal was analysed in the identical way as the pulsar observations. Magnitude limits are 90% confidence upper limits (lower limits on the magnitude scale).

The radio pulses have more power at the second harmonic than at the fundamental frequency, so it is notable that there are peaks at the second harmonic in most of the data. Nevertheless, even the highest has 3% probability of being noise and the next highest has 7% probability⁸. There are not even noise peaks at the fundamental frequency.

We have no compelling evidence of pulsations from the source. Conservative upper limits for 100% sinusoidal modulation are 24 mag in the band of 0.7–0.9 μ m and 16.4 mag in the band of 1.4–2.3 μ m. These limits correspond to about a 3% modulation of the total light from the red candidate or a variability of <0.03 mag.

We suspect that the red candidate is a reddened ordinary star falling near the position of the pulsar by chance. The commonest stars visible at great distances are K giants; their intrinsic $J-K$ colour is 0.7 mag and their $R-K$ colour is 1.8 mag (ref. 9). If the red candidate is a K giant, $A_v = 8$ from comparison of observed and intrinsic colours. The K magnitude of the star is 11.5 after correction for extinction with $A_K = 0.8$. A typical M_K for K giants is -2.5 , which at 6.3 kpc gives the observed m_K . The density of K giants¹⁰ in the direction of the pulsar is 0.6 to 2.5×10^{-4} pc⁻³. If all objects within 4 arc s of the radio position are studied for connection with the pulsar, there is a cumulative probability of 1–3% that a K giant of at least the required flux would be found by chance.

The two-filter CO photometric index between 2.2 and 2.33 μ m is defined as follows. It is zero magnitude for unreddened A0 stars with a Rayleigh-Jeans spectrum, and the index increases as the spectrum becomes steeper than that of an A0 star. For a source as red as the proposed counterpart to PSR1937+214, a negative index would be expected, whereas our measured value was $+0.19 \pm 0.03$. This value is a strong

indication for the presence of CO absorption bands of a strength typical for K giants⁶. We conclude that the red candidate is probably a red giant.

We thank Jerry Hudson for calculating the pulsar ephemeris used for real time averaging and S. Djorgovski for prepublication results and many discussions. These observations were supported by the NSF under grant AST 81-14717 and performed in part under the auspices of the Director, Office of Energy Research, Office of High Energy and Nuclear Physics, Division of High Energy Physics, Department of Energy under contract no. DE-AC03-76SF0098.

Received 7 June; accepted 1 September 1983.

1. Backer, D., Kulkarni, S., Heiles, C., Davis, M. & Goss, W. *IAU Circ. No. 3743* (1982).
2. Backer, D., Kulkarni, S., Heiles, C., Davis, M. & Goss, W. *Nature* **300**, 615-618 (1982).
3. Djorgovski, S. *Nature* **300**, 618-620 (1982).
4. Manchester, R. M., Peterson, B. A. & Wallace, P. T. *IAU Circ. No. 3795* (1983).
5. Lebofsky, M. & Rieke, G. *IAU Circ. No. 3809* (1983).
6. Frogel, J. A., Persson, S. E., Aaronson, M. & Matthews, K. *Astrophys. J.* **220**, 75-97 (1978).
7. Middleditch, J. *Astrophys. J. Lett.* **257**, L71-L75 (1982).
8. Groth, E. *Astrophys. J. Suppl.* **29**, 285-302 (1975).
9. Johnson, H. A. *Rev. Astr. Astrophys.* **4**, 193-206 (1966).
10. Wrاندemark, S. *Astr. Astrophys.* **58**, 99-103 (1977).
11. Becker, R. & Helfand, D. *Nature* **302**, 688-690 (1983).

A unified structural approach to linear carbon polytypes

R. B. Heimann*, J. Kleiman & N. M. Salansky

3M Canada Inc., Institute for Aerospace Studies,
University of Toronto, Downsview, Ontario, Canada M3H 5T6

Compounds obtained by high-temperature treatment of graphite are thought to be linear carbon polytypes with sp-configuration and carbon-carbon chains, either conjugated triple bonded or cumulated double bonded, parallel to the c-axis. At least 10 crystallographically documented linear carbon forms (carbynes) have been reported. To relate the contradictory and confusing information found in the literature, a simple classification is suggested here based on a linear relationship between number of atoms in a chain (n) and the unit cell parameters a_0 and c_0 of all carbyne forms known. It is assumed that chains are kinked, and that the distribution of kinked spacings may be a function of the temperature of formation of the carbyne forms in question.

Investigations of the carbynes were triggered by their possible utilization as nose tips for space re-entry vehicles¹, as potential material for ultrastrong fibres^{2,3}, and as superconductors at ambient temperatures⁴.

Twenty years ago, a poorly crystallized compound was obtained⁵ by oxidative dehydrocondensation of acetylene, using copper catalyst. The X-ray diffraction pattern suggested the presence of densely packed, rectilinear, conjugated triple-bonded carbon ('polyne') chains⁶. Prolonged heating at 1,000 °C under reduced pressure yielded a fully crystalline product, which was identified as a third allotropic modification of carbon and named 'carbyne'⁷. Electron microdiffraction later revealed the presence of two phases, one with polyne-type bonds (α -carbyne) and one with cumulene-type bonds (cumulated double bonds; β -carbyne)⁸.

During the next few years, carbyne-related compounds were observed by the condensation of carbon vapour⁹, as products of regraphitization of diamond¹⁰, by the interaction of shock waves with carbon-bearing material¹¹, and by the heating of graphite to temperatures exceeding 2,600 K (refs 12, 13). Electron diffraction and X-ray diffraction (EDX) investigations yielded several new members of the carbyne family: carbon VI (ref. 14), carbon VIII (ref. 15), carbon IX (ref. 13), carbon

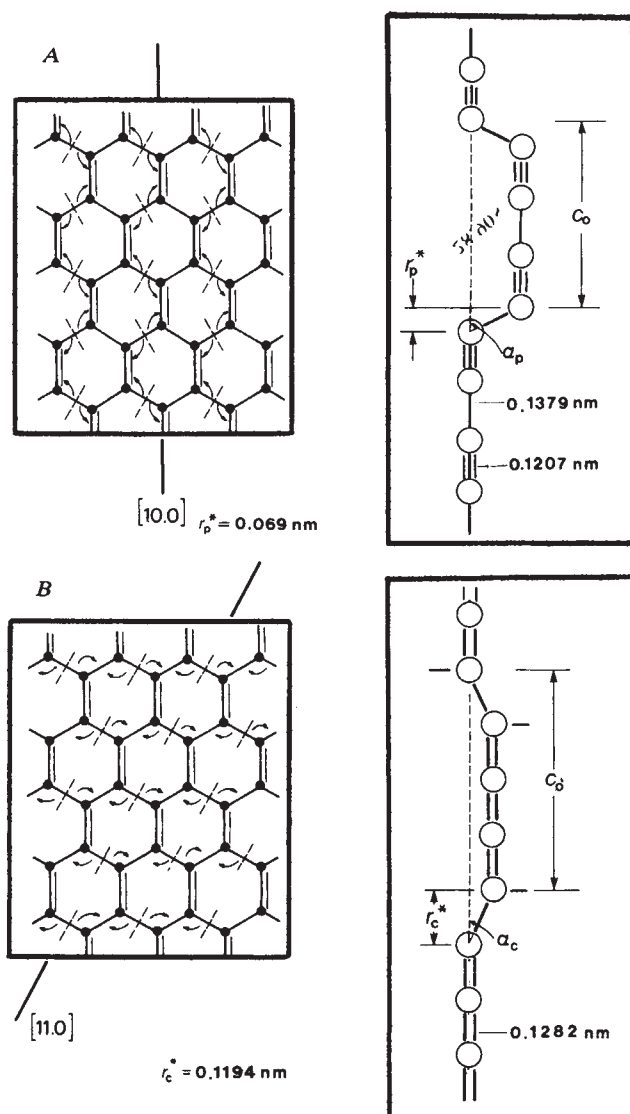


Fig. 1 Conformation of kinked chains in the basal plane of graphite. A, Conjugated triple-bond (polyne) chains parallel [10.0] according to Whittaker³³. B, Cumulated double-bond (cumulene) chains parallel [11.0] according to Heimann *et al.*¹⁷.

X-XIII (A. G. Whittaker, personal communication and ref. 16), and a new form, carbon XIV, presumably obtained by regraphitization of shock-formed diamond¹⁷. Moreover, a natural carbon mineral was found in shock-fused graphite gneisses in the Ries Crater, Bavaria, and named 'chaoite'¹⁸.

To relate the various carbyne forms by a simple classification scheme based on chain lengths we conjectured that: (1) carbynes are formed by a crystallographic arrangement of C-C chains with either conjugated triple or cumulated double bonds; and (2) c-axes lengths are functions of the number of carbon atoms, n , in the chain.

Doubts about the existence of any carbyne form at all have been expressed by Smith and Buseck¹⁹, Lumpkin²⁰ and Bundy²¹. The former authors regard all reports on carbynes as misinterpretations of electron diffraction patterns which in their opinion always belong to sheet silicates. Whittaker¹³ showed, however, that, on heat treatment, highly purified carbon (impurity content 6 p.p.m.) yielded carbon IX. Moreover, EDX analyses with a scanning transmission electron microscope (STEM) of a chaoite-like carbon phase obtained by shock-loading of graphite powder showed the absence of any detectable contamination¹⁷.

Evidence for the existence of C-C chains in carbynes is found in an extensive body of literature dealing with X-ray diffraction^{5,6} and various spectroscopic methods, such as IR spectroscopy^{22,23}, negative ion mass spectroscopy²⁴, laser

* Present address: Atomic Energy of Canada Ltd, Whiteshell Nuclear Research Establishment, Pinawa, Manitoba, Canada.

Table 1 Crystallographic data of carbyne forms

	β -Carbyne (ref. 29)	Carbon-XIV (ref. 17)	Carbon-VI (ref. 14)	Carbon-IX (ref. 13)	Chaoite (ref. 18)	Carbon-VIII (ref. 15)	α -Carbyne (ref. 29)
a_0 (nm)	0.824	0.870	0.923*	0.944*	0.8948	0.9098	0.894
c_0 (nm)	0.768	0.956	1.224	1.250	1.4078	1.482	1.536
a_0/c_0	1.073	0.910	0.754	0.755	0.635	0.614	0.581
$\frac{1}{3}V$ (nm ³)	0.451	0.626	0.903	0.965	0.976	1.062	1.058
Z	72	—	—	—	168?	—	144
ρ (g cm ⁻³)	3.13	—	2.9	—	3.43	—	2.68
n	6	8†	10†	10†	11†	12†	12
Space group	P3 _{1,2}	—	P3 or P3 _{1,2}	Rhombohedral (R3?)	P6/mmm? P3 or P3 _{1,2}	—	P3 _{1,2}

* a_0 Given for a hexagonal cell ($a_{\text{hex}} = a_{\text{rhom}}\sqrt{3}$).

† Assumed values of n .

Raman spectroscopy²⁵, and X-ray photoelectron spectroscopy²⁶. The chains are thought to be arranged parallel to the c -axes, and bonding is achieved by intermolecular linkage. The interchain distance of 0.293 nm (ref. 27) is sufficiently small to permit the interaction of unoccupied π -orbitals in a bonding mode²⁸. Sladkov and Kudryavtsev²⁷ assigned a tentative unit cell with three chains per cell which obeys the space group symmetry P3 or P3_{1,2}. The space groups P3 or P3_{1,2} were assumed by Whittaker for carbon VI (ref. 14). Carbon IX is thought to possess a rhombohedral space group¹³. Thus, the point symmetry group of carbynes may be 3 which contains the space groups P3, P3₁, P3₂ and R3. Table 1 shows crystallographic data of carbyne forms reported in the literature. The n values refer to the number of carbon atoms in the chains, and were given as 6 for β -carbyne and 12 for α -carbyne. We have assumed the other n values. Clearly, a discrete identity period cannot be assigned in the c -direction for infinite and rectilinear C—C chains. To resolve this problem, C—C chains can be divided, in principle, into three types: (1) n -membered C—C chains with a helical structure whose threads with a pitch $\tau = 360^\circ/n$ generate a lattice translation analogous to a crystallographic n -fold screw axis; (2) zig-zag chains; and (3) straight portions of chains separated by kinks, which constitute a 2-fold screw axis. Chains of the first type are present in polyacrylonitrile³⁰; chains of the second type are present in trans-polyalkenamers³¹; chains of the third type have been identified in crystalline polyethylene³². Polyene and cumulene chains are very rigid owing to the π -component in the linear molecular-orbital bonding system, and are, therefore, unlikely to form a helix. Kinked chains, however, can be formed through a mechanism first introduced by Whittaker³³ to explain the formation of polyene chains, by the breaking of single bonds and the shifting of electrons towards adjacent double bonds in a graphite basal plane (see Fig. 1a). An analogous mechanism has been conjectured for the formation of cumulene bonds¹⁷ and is shown in Fig. 1b. Moreover, kinked chains provide free bonds by splitting double bonds. Splitting of triple bonds leads likewise to free bonds adjacent to kink sites but requires a reorganization of the remainder of the bonds thus switching the structure to a cumulene-type. These free bonds are available for bonding stabilizing bulky alkyl or aryl groups, which prevent lattice collapse by keeping sufficient distances between the chains³⁴; adding stabilizing copper⁶ or silicon atoms (J.K. *et al.*, work in preparation); or cross-linking carbon chains³³, which may explain the exceptional hardness and thermal stability of some carbynes^{12,14}. Therefore, it is reasonable to assume that the cumulene-type carbynes with an intermolecular linkage between the chains due to the free bonds will be more stable than the polyene-type carbynes. The kink spacings define the c_0 -axes lengths as

$$c_0(\text{polyene}) = \frac{n}{2} r(\text{C}\equiv\text{C}) + \left(\frac{n}{2} - 1\right) r(\text{C}=\text{C}) + r_p^* \quad (1)$$

and

$$c_0(\text{cumulene}) = (n-1)r(\text{C}=\text{C}) + r_c^* \quad (2)$$

where

$$r_p^* = r(\text{C}=\text{C}) \cos 60^\circ, \quad r_c^* = r(\text{C}=\text{C}) \cos 30^\circ$$

The values $r(\text{C}=\text{C})$, $r(\text{C}\equiv\text{C})$, $r(\text{C}\equiv\text{C})$ are 0.1379 nm, 0.1282 nm and 0.1207 nm, respectively, and were obtained from Stoeckert's relation³⁵. This more intuitive model is corroborated by approaches given by the Pariser-Parr-Pople theory³⁶, which predicts C—C bond distances in terms of the π -bond order³⁷. It has been suggested¹⁷ that the kink angles, α_p of 60° for a polyene chain and α_c of 30° for a cumulene chain, are remainders of the original bonding angle in the basal plane of graphite. This assumption, however, meets intrinsic difficulties in explaining the mode of kink generation. Alternatively, a tetrahedron model should be considered based on the results obtained by Natta *et al.*³¹ on linear *trans*-polyalkenamers. The first member of the series is polyacetylene, the last is polyethylene. Members of these homologous series consist of planar zig-zag chains whose lengths are shortened on the basis of the principles of staggered bonds extended to the double bonds³⁸. According to this principle, twisting of the n -membered chain ($n > 2$) around the single bond adjacent to each double bond produced kinks acting as 2-fold screw axes. The identity period of an even *trans*-polyalkenamer can be calculated according to the equation³¹

$$c_0 = (n-1)r(\text{C}=\text{C}) \sin(\phi_1/2) + r(\text{C}\equiv\text{C}) \sin(\phi_2 - \phi_1/2) \quad (3)$$

where $r(\text{C}=\text{C}) = 0.154$ nm, $r(\text{C}\equiv\text{C}) = 0.134$ nm, $\phi_1 = 110^\circ$, $\phi_2 = 120^\circ$, and n number of carbon atoms in the repeating unit. The term $\sin(\phi_2 - \phi_1/2)$ represents the torsional angle of 65° .

If we think of a carbyne as a fully dehydrogenated polyalkenamer, the single bonds connecting the methylene groups are converted into either double or triple bonds, and the chemical repeating units are separated by kinks whose original double bonds are split providing the radical centres with sp^2 configuration³⁴ mentioned above. The equation for the c_0 distances (equation (3)) is then transformed into

$$c_0 = (n-1)r(\text{C}=\text{C}) \sin(\phi_1^*/2) + r(\text{C}=\text{C}) \sin(\phi_2 - \phi_1/2) \quad (4)$$

where $\phi_1^* = 180^\circ$ as required for a strictly rectilinear macromolecular chain of carbon atoms in sp state, that is the zig-zag chain configuration of the polyalkenamer is lost by dehydrogenation. Equation (4) is identical to the empirically-derived equation (2). Table 2 gives a comparison of observed (Table 1) and calculated c_0 -axes lengths for carbynes with cumulene (C)- and polyene (P)-type chains. From the calculated kink angle that gives ideal fit, one obtains the narrow ranges

$$23^\circ < \alpha_c < 25^\circ \text{ for cumulene chains}$$

and

$$60^\circ < \alpha_p < 65^\circ \text{ for polyene chains}$$

Table 2 also gives observed c_0 distances for *trans*-polyoctenamer, *trans*-polydecenamer, and *trans*-polydodecenamer, as well as the calculated values according to a modified equation (3)³¹. The last column of Table 2 shows the c_0 -distances calculated for fully dehydrogenated polyalkenamers according to

Table 2 Comparison of calculated and observed c_0 lengths of different carbyne forms

Phase	n^*	c_0 (obs.) (nm)	c_0 (calc.) [†] (nm)	$\alpha_{\text{theor}}^{\ddagger}$ (deg)	Bond type	Polyalkenamer§		Dehydrogenated polyalkenamer
						c_0 (obs.) (nm)	c_0 (calc.) (nm)	c_0 (calc.) (nm)
β -carbyne	6	0.768	0.760	22.93	C	—	0.742	0.766
Carbon XIV	8	0.956	0.965	64.44	P	0.990	0.991	1.022
Carbon VI	10	1.224	1.224	60.00	P			
Carbon IX	10	1.250	1.224	46.51	P	1.240	1.240	1.278
			1.273	45.76	C			
					C			
Chaoite	11	1.408	1.401	23.97	C			1.407
Carbon VIII	12	1.482	1.483	60.21	P			
α -carbyne	12	1.536	1.530	24.18	C	1.485	1.491	1.535

* Number of C atoms in the chain.

[†] According to equations (1) and (2).[‡] Kink angle required to match observed and calculated c_0 axes lengths exactly.

§ Ref. 31.

|| Equation (4) with shortened bonds according to Stoicheff's rule³⁵.

equation (4) using shortened bond distances obtained by Stoicheff's rule³⁵. The c_0 values, in fact, are close to the ones observed for the carbyne forms. The fit for carbon XIV is not very good, probably due to its assumed polyene-type chain. Note that the range of the torsional kink angles $23^\circ < \alpha_c < 25^\circ$ corresponds almost exactly to the angle obtained for the polyalkenamer structure and to its complementary angle to 90° , respectively. Figure 2a shows the linear relationship between n and the observed c_0 parameters. The dashed line relates to the polyene-type bond with $\alpha = 60^\circ$ and obeys the equation $c_0 = 0.1295n - 0.071$; the unbroken line relates to the cumulene-type bond with $\alpha = 30^\circ$ and yields $c_0 = 0.1283n - 0.010$. The fit of the data points to the theoretical function is excellent as mirrored by the regression coefficients of 0.997 for the polyene-type bond and 0.999 for the cumulene-type bond, respectively.

Figure 2b gives the a_0 parameters of several carbyne forms as a function of n . The a_0 values of carbon X–XIII were taken from d_{max} values on a c -axis pattern as obtained from electron diffraction data by Whittaker *et al.*¹⁶. Assignment of n values to these compounds is only tentative. The shallower slope of the cumulene-type line, as opposed to the polyene-type line, may reflect a lattice contraction in the a direction owing to increasing cross-bonding^{22,23} between the kinked chains.

Plots of the cell volumes, V and the axial ratios, a_0/c_0 against n yield linear relationships for all carbyne forms and can be expressed as $V = 0.1407n - 0.1696$ ($r^2 = 0.982$) and $a_0/c_0 = -0.081n + 1.555$ ($r^2 = -0.995$).

It is reasonable to assume that the chain lengths are functions of the temperature of formation of the carbynes, that is higher temperatures favour cracking of the chains⁶. Experimental results by Whittaker³³ suggest that the carbyne form which is stable slightly above 2,600 K is chaoite ($n = 11$) but that the carbyne form which melts around 3,800 K to give liquid carbon with optical properties of a carbyne is β -carbyne ($n = 6$)³⁹. This concept of a temperature-dependent 'equilibrium' of chain lengths is consistent with the exceptionally narrow stability ranges of the individual carbyne forms of only ~ 200 K (ref. 33). Note that a random distribution of the chain lengths around a mean value characteristic for each particular temperature of formation might take place.

Received 28 April; accepted 2 September 1983.

- Whittaker, A. G. & Evangelides, J. S. *Carbon* **15**, 435 (1977).
- Perelkin, K. E., Korshak, V. V. & Kasatichkin, V. I. *Dokl. Akad. Nauk. SSSR* **220**, 1376–1379 (1975).
- Whittaker, A. G. US Patent 4,248,909 (1981).
- Little, W. A. *Phys. Rev.* **134**, 1416–1424 (1964).
- Korshak, V. V. *et al. Dokl. Akad. Nauk SSSR* **136**, 1342–1344 (1961).
- Sladkov, A. M. *Soviet Sci. Rev. B3*, 75–110 (1981).
- Sladkov, A. M., Kasatichkin, V. I., Kudryavtsev, Yu. P. & Korshak, V. V. *Izv. Akad. Nauk-SSSR, Ser. Khim.* **12**, 2697–2704 (1968).
- Kasatichkin, V. I., Popov, N. M., Sladkov, A. M., Kudryavtsev, Yu. P. & Korshak, V. V. *Dokl. Akad. Nauk SSSR* **177**, 358–360 (1967).
- Kasatichkin, V. I., Savranski, V. V., Smirnov, B. N. & Melnichenko, V. M. *Dokl. Akad. Nauk SSSR* **217**, 796–799 (1974).
- Hall, H. T. *Science* **192**, 868 (1970).
- Setaka, N. & Sekikawa, Y. *J. Am. ceram. Soc.* **63**, 238–239 (1980).
- Nelson, L. S., Whittaker, A. G. & Tooper, B. *High Temp. Sci.* **4**, 445–477 (1972).
- Whittaker, A. G., Neudorffer, M. E. & Watts, E. J. *Carbon* (in the press).
- Whittaker, A. G. & Wolten, G. M. *Science* **178**, 54–56 (1972).
- Whittaker, A. G. *Carbon* **17**, 21–24 (1979).
- Whittaker, A. G., Watts, E. J., Lewis, R. S. & Anders, E. *Science* **209**, 1512–1514 (1980).
- Heimann, R. B., Kleiman, J. & Salansky, N. M. *Carbon* (in the press).
- El Goresy, A. & Donnay, G. *Science* **161**, 363–364 (1968).
- Smith, P. P. K. & Buseck, P. R. *Proc. Lunar planet. Sci.* **12B**, 1167–1175 (1981); *Science* **216**, 984–986 (1982).
- Lumpkin, G. R. *Lunar planet. Sci.* **12**, 631–633 (1981).
- Bundy, F. P. *J. geophys. Res.* **85**, 6930–6936 (1980).
- Kasatichkin, V. I. *et al. Dokl. Akad. Nauk SSSR* **153**, 346–349 (1963).
- Hay, A. S. *J. polym. Sci. A1*, **V7**, 1625–1634 (1969).
- Stuckey, W. K. & Whittaker, A. G. *Abstr. 10th Conf. Carbon*, Lehigh University TP-177, 278–279 (1971).
- Nakamizo, M., Kammereck, R. & Walker, P. L. *Carbon* **12**, 259–267 (1974).
- Sergushin, I. P. *et al. Zh. strukt. Khim.* **18**, 698–700 (1977).
- Sladkov, A. M. & Kudryavtsev, Yu. P. *Priroda* **58**, 37–44 (1969).
- Patai, S. (ed.) *The Chemistry of the Carbon-Carbon Triple Bond* (Wiley, New York, 1978).

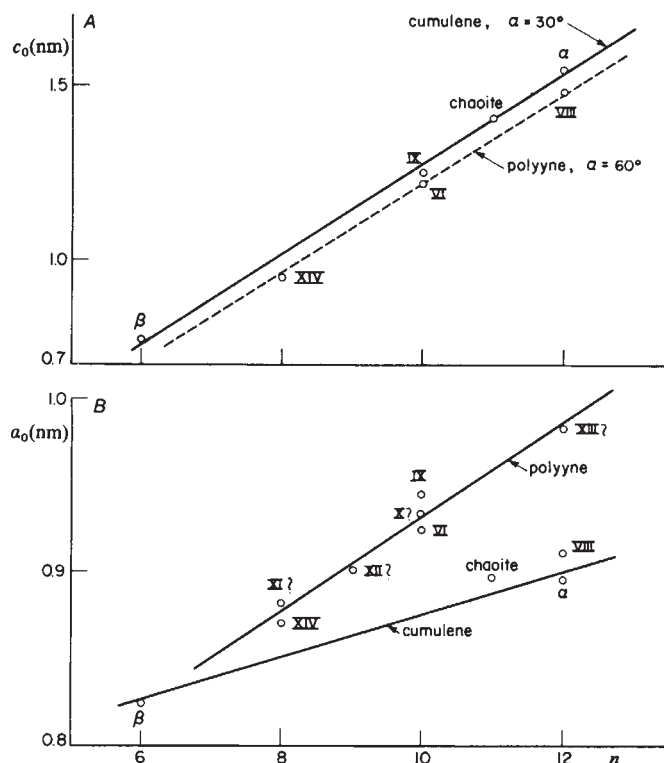


Fig. 2 Relationship between the number of atoms in the chain, n , and A , the c_0 parameter for polyene ($\alpha = 60^\circ$) and cumulene ($\alpha = 30^\circ$) arrangement; B , the a_0 parameter.

29. Kasatochkin, V. I., Korshak, V. V., Kudryavtsev, Yu. P., Sladkov, A. M. & Sterenberg, L. E. *Carbon* **11**, 70–72 (1973).
30. Lindenmayer, P. H. & Hosemann, R. *J. appl. Phys.* **34**, 42–45 (1963).
31. Natta, G., Bassi, I. W. & Fagherazzi, G. *Eur. Polym. J.* **5**, 239–260 (1969).
32. Reneker, D. H. *J. polym. Sci.* **59**, S39–S42 (1962).
33. Whittaker, A. G. *Science* **200**, 763–764 (1978).
34. Jansta, J. & Dousek, F. P. *Carbon* **18**, 433–437 (1980).
35. Stoicheff, B. P. *Tetrahedron* **17**, 135–145 (1962).
36. Julg, A. *J. chem. Phys.* **65**, 541–548 (1968).
37. Coulson, C. A. *Valence* (Clarendon, Oxford, 1952).
38. Perego, G. & Bassi, I. W. *Makromolek. Chem.* **61**, 198–208 (1963).
39. Whittaker, A. G. & Kintner, P. L. *Carbon Vapour Pressure in the Range 3450 to 4500 K and Evidence for Melting at ~3800 K* (Rep. Aerospace Corp., Materials Science Lab., El Segundo, 1981).

Bouncing internal bores of Ardmucknish Bay, Scotland

S. A. Thorpe & A. J. Hall

Institute of Oceanographic Sciences, Wormley, Godalming, Surrey GU8 5UB, UK

S. Hunt

Department of Physics, University of Surrey, Guildford, Surrey GU2 5XH, UK

Propagating fronts and internal waves have been observed in the sea using bottom-mounted side-scan sonar¹ in Ardmucknish Bay, north of Oban on the west coast of Scotland. The fronts, between brackish water and seawater, are generated when less saline, and so less dense, water discharges tidally into the sea from the mouth of a sea loch. The fronts appear to propagate from the direction of the loch, as might be expected, but the internal waves were found to propagate almost at right angles to this direction. A time-lapse cine camera was mounted on a nearby hill and was used to investigate the phenomena, which are visible on the surface because of their effect on surface waves and hence on the brightness of the sea surface. The internal waves are found to be generated when the fronts reflect from the coastline on one side of the bay, and this explains their direction of propagation. The movement of fronts influence the horizontal transfer (of say nutrients) from the sea loch into the surrounding coastal water.

The source of the fronts is the tidal discharge into the sea of brackish water from Loch Etive through the Falls of Lora, a shallow sill at the mouth of the sea loch (see Fig. 1a). Remarkable in these observations was that whilst the fronts generally propagated past the sonar in a westerly or southwesterly direction, away from the falls, the groups of 3–7 internal waves propagated to the south or south-east. J. E. Simpson sent us a photograph taken in September 1982 from a nearby hill, Beinn Lora, which shows 5 or 6 bands on the sea surface at position A in Fig. 1a, tentatively identified as internal waves with wavelength ~52 m. Their position and curvature suggests that they too were propagating from the north-west. Here we offer an explanation of these observations and present further measurements of the fronts, including some made by a towed linear array of thermistors.

From 18 to 21 April 1983 a 16-mm cine camera taking photographs every 30 s was mounted at position 0 in Fig. 1b, 275 m above sea level on Beinn Lora, overlooking Ardmucknish Bay. Visual observations were also made which usefully supplemented the photographic record. Fronts were seen emerging from the passes around Eilean Mor some 2 h after high water at Oban (high water is 0–20 min later at the falls), and the film shows their spreading across Ardmucknish Bay. Figure 1b (at A) shows the positions of a front at 20-min intervals on 18 April, when winds were light from the north. The front appeared as a single continuous dark curved line on the water surface. The presence of fronts and subsurface internal waves can often be detected at the surface because they induce changes in near-

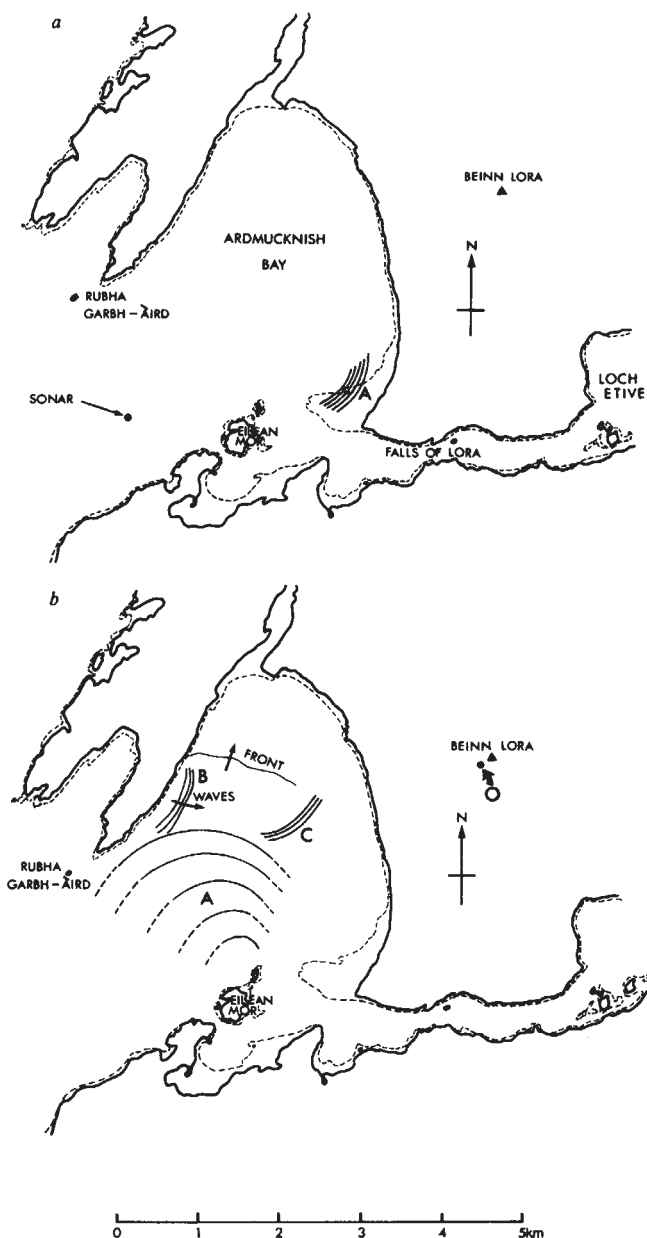


Fig. 1 a, Chart showing Ardmucknish Bay, the position of the sonar and the internal waves (at A). Mean high and low water levels are shown by full and dashed lines respectively. b, As in a, but showing: A, the position of front on 18 April 1983 at 20-min intervals; B, the reflection of the front; C, the internal waves propagating to the south-east. 0 is the camera position on Beinn Lora.

surface currents which in turn modify the surface waves which scatter light from the sky to the observer^{2,3}. The front advanced at a speed which decreased from 0.43 ± 0.04 to $0.35 \pm 0.15 \text{ m s}^{-1}$ (one standard deviation is indicated) as it crossed the bay. On arrival at the far side of the bay the front was reflected as sketched at B in Fig. 1b, developing a train of waves which travelled back across the bay. Their position 6 h 55 min after high water at Oban is shown at C in Fig. 1b when three waves were visible. The wavelength of the waves was about 64 m and they travelled at $0.29 \pm 0.06 \text{ m s}^{-1}$. Around Rubha Garbh-àird a complex pattern of surface markings could be seen, but was not resolved on the film. Reflection of the groups of waves from the shoreline at the eastern side of the bay could not be detected.

In March–April 1982 currents were measured by a moored Aanderaa current meter at a depth of 22 m near the position of the sonar where the water depth is 35 m. During the first 7 h after high water, the mean M_2 tidal currents are $<0.05 \text{ m s}^{-1}$,

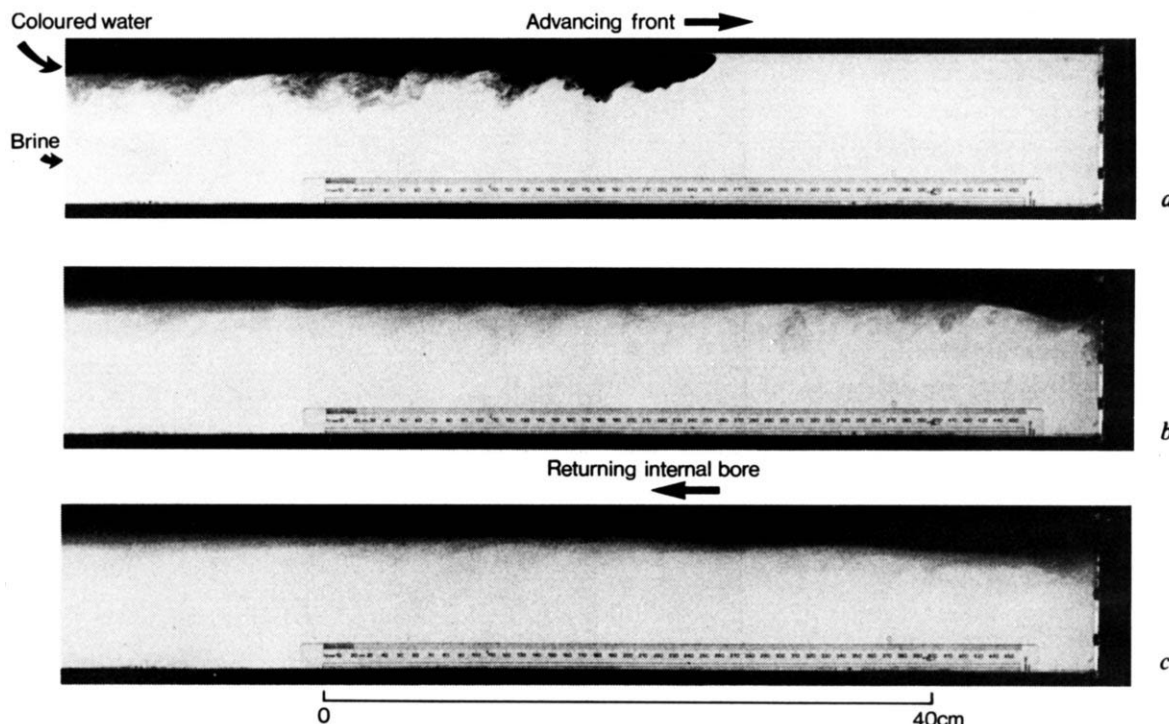


Fig. 2 A front, or gravity current, reflecting at the end of a closed tube in the laboratory. The tube 180 cm long, 10 cm high and 3 cm wide, was completely filled in a vertical position with layers of clear brine and water (10 cm deep and dyed with potassium permanganate) with density difference 0.072 g cm^{-3} , and tilted to the horizontal a few seconds before *a*, which shows the spreading gravity current head. This meets the end of the tube (at *b*) and reflects back (*c*), producing a train of internal waves (the first is just visible).

predominantly east-north-east or southwesterly in direction, and are much less than the speeds of the front or waves.

A simple laboratory experiment with two layers, brine and coloured water filling a tube, illustrates some of the processes involved (Fig. 2). The advancing front (Fig. 2*a*) has the form of a gravity current⁴. The gravity current meets the end wall of the tube (Fig. 2*b*), mimicking the arrival of the front at the shoreline on the north-west side of the bay, and produces a depression on the interface. This, in recovering under buoyancy forces, spreads as an internal bore or soliton, propagating back into the now stratified fluid (Fig. 2*c*) and developing internal waves as described elsewhere⁵⁻⁷.

A front was again clearly visible on 21 April advancing across Ardmucknish Bay at a mean speed of $0.27 \pm 0.06 \text{ m s}^{-1}$. Winds on this occasion were westerly, about 5 m s^{-1} . Temperature sections were made through the front by an array of thermistors mounted on a streamlined spar hanging below a catamaran towed at $\sim 1.6 \text{ m s}^{-1}$ by RV *Calanus*. Data were recorded at a frequency of 4.5 Hz, temperature being resolved to better than

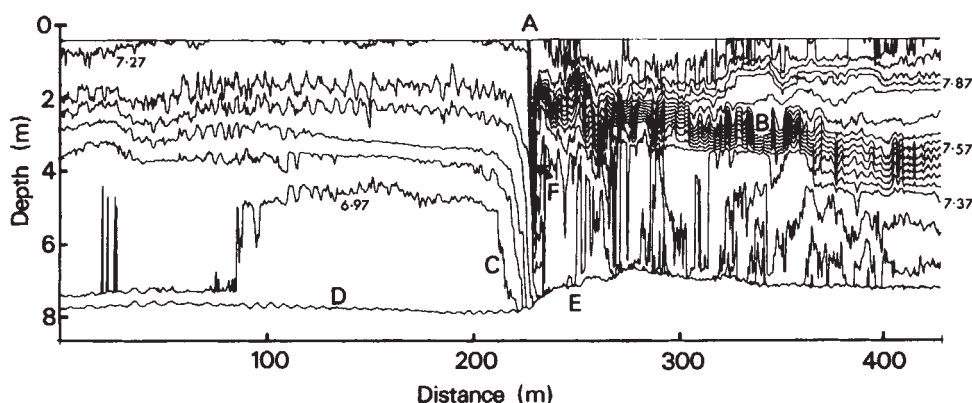
1 mK. As shown in Fig. 3 the front is remarkably narrow with intense gradients; at 1 m depth a change of 0.36 K occurs in a horizontal distance of 0.35 m and the total temperature jump across the front, $\sim 0.55 \text{ K}$, occurs in a distance of $< 1 \text{ m}$. A change in the component of current parallel to the front was made apparent by the rapid tilting of the spar supporting the thermistors as the front was crossed (see legend to Fig. 3), but the shear could not be estimated directly. Differences of 0.16 m s^{-1} have been observed at the sonar (Fig. 1*a*)¹. The brackish layer of water flowing out of Loch Etive appears to be $\sim 3.5 \text{ m}$ in depth, much less than the total water depth of 30–40 m in the central part of the bay.

Supposing that the front advances as a gravity current at a speed u given by⁴

$$u^2 = 2gh \left(\frac{\rho_2 - \rho_1}{\rho_1} \right) \quad (1)$$

where g is the acceleration due to gravity, h is the depth of the brackish layer of density ρ_1 , and ρ_2 is the density of the seawater,

Fig. 3 Contours of temperature, in $^{\circ}\text{C}$, at 60-mK intervals, through a front from the seaward side into the brackish layer. The front, advancing to the left, is marked at A. The horizontal scale is converted to distance using the ship's speed over the seabed. Water is thermally stratified on both sides of the front. The intense gradients at B appear to mark the bottom of the density flow. The contours are slightly depressed at C before the front arrives. The line D marks the depth of the lowest thermistor on the towed spar. This rises at E due to the increased current as the front is crossed. The isotherms at F immediately behind the front at the head of the gravity current are deeper than at B (compare with Fig. 2*a*).



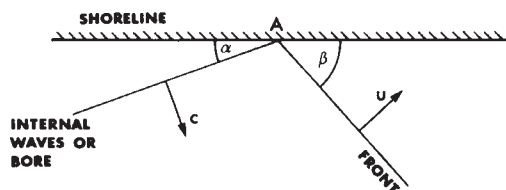


Fig. 4 Reflection of a front at a vertical shoreline. The speed of point A along the shoreline is $u/\sin \beta$ or $c/\sin \alpha$.

we can estimate the fractional density difference,

$$\frac{\rho_2 - \rho_1}{\rho_1} = 1.1 \times 10^{-3}$$

This is much greater than could be produced by the observed temperature change (Fig. 3), and is consistent with a salinity change of ~1%.

The speed c of the internal waves is given approximately⁸ by

$$c^2 = gh \left(\frac{\rho_2 - \rho_1}{\rho_1} \right) \quad (2)$$

which by equation (1) $= u^2/2$ and this is reasonably consistent with the observed relative values of c and u . The relation $c = u/\sqrt{2}$ implies that the angles which the lines of the waves and front make with shoreline, α and β respectively (Fig. 4), are such that

$$\frac{\sin \beta}{\sin \alpha} = \sqrt{2} \quad (3)$$

This is roughly what was observed. The value of α must be $<45^\circ$.

This estimate is, however, based on a simplified conception of the dynamics and can give, at best, only an approximate prediction. The internal bore reflecting from the shoreline is of finite amplitude, even when it degenerates into a packet of internal waves. Particle speeds exceeding 50% of the internal wave phase speed have been measured by the sonar¹, and thus equation (2) is inadequate⁸. Equation (2) also fails to account for the flow behind the gravity current, which induces a vertical shear in the stratified layer into which the bore propagates. Moreover the bore must be such as to reduce to zero the current component normal to the shoreline. There is evidence⁹ that the factor 2 appearing in equation (1) may be too great. The finite depth of the lower layer may also be taken into account. Laboratory experiments are being made to study reflection in more detail.

Nevertheless, given the coarse estimates which we have made, the directions of propagation of the internal waves at the sonar site and those reported by Simpson are explained. They result from the reflection, or 'bouncing', of a front at the shore line particularly at the north-west side of Ardmucknish Bay, and a transition in character to an internal wave packet as the pulse returns into the stratified environment produced behind the spreading front. There remains, however, one difficulty. Simpson's waves were seen only 1 h after high water at Oban. To reconcile this time with our observations and simple interpretation, the propagation speeds must have been lower than we found so that the waves reached position A in Fig. 1a some 11 h after the front generating them passed Eilean Mor. With variable winds, tides and outflow from Loch Etive the movement of fronts and waves around Ardmucknish Bay is probably sometimes even more complicated than we observed. The fronts affect the extent of the exchange between the sea loch and the surrounding coastal water. The processes which influence the speed and decay or disintegration of the fronts are thus important in influencing transport of material (such as nutrients) in their passage from land to sea.

We acknowledge the help and enthusiasm of the master and crew of RV *Calanus* of the Scottish Marine Biological

Association, and of Mr John Hibbert who assisted with the photography. We thank Dr J. E. Simpson for the photograph and information.

Received 13 June; accepted 9 September 1983.

1. Thorpe, S. A. & Hall, A. J. *Continental Shelf Res.* **1**, 353-384 (1983).
2. Hughes, B. & Gower, J. F. R. *J. geophys. Res.* **88**, 1809-1824 (1983).
3. Apel, J. R., Byrne, H. M., Proni, J. R. & Charnell, R. L. *J. geophys. Res.* **80**, 865-881 (1975).
4. Benjamin, T. B. *J. Fluid Mech.* **31**, 209-248 (1968).
5. Maxworthy, T. *J. Fluid Mech.* **96**, 47-64 (1980).
6. Djordjevic, V. D. & Redekopp, L. G. *J. phys. Oceanogr.* **8**, 1016-1024 (1978).
7. Segur, H. & Hammack, J. L. *J. Fluid Mech.* **118**, 285-304 (1982).
8. Benjamin, T. B. *J. Fluid Mech.* **29**, 559-92 (1967).
9. Simpson, J. E. & Britter, R. E. *J. Fluid Mech.* **94**, 477-495 (1979).

Continental freeboard, sedimentation rates and growth of continental crust

Scott M. McLennan & S. R. Taylor

Research School of Earth Sciences, Australian National University, Canberra, ACT 2601, Australia

To understand how the Earth has evolved requires a model for the timing and nature of the formation of the continental crust. Of the two contrasting models, the first and most popular is that the continental crust of the Earth has grown throughout geological time. Models of linear or exponential increasing growth with time^{1,2} are not currently accepted, and it is now generally held that much of the crust was in place by 2,500 Myr with a particularly active period of crustal growth and differentiation between about 3,200 and 2,500 Myr (refs 3-13). The second model argues that the present mass of the continental crust formed very early in Earth history ($>4,000$ Myr) and has subsequently been recycled through the mantle in a steady-state fashion such that the mass of the crust has not changed during most of Earth history¹⁴⁻¹⁶. Resolving these conflicting views is not simple because interpretation of the isotopic evidence is non-unique. The resolution of this debate lies in the interpretation of the record of continental freeboard¹⁷⁻¹⁹ and in the plausibility of large scale crustal recycling through the mantle via sediment subduction. We point out here that although the record of continental freeboard is consistent with either limited growth or no-growth models, recent sedimentation rates appear insufficient to support no-growth models. Note that evidence against a no-growth model is not necessarily evidence against minor sediment subduction.

It has been demonstrated that the isotopic record is consistent with no-growth crustal recycling models¹⁴⁻¹⁶. Early modelling¹⁴ indicated that at present, $\sim 5 \text{ km}^3 \text{ yr}^{-1}$ of crust would have to be recycled through the mantle and that this rate would have to increase exponentially with geological age to sustain a no-growth model which predicts the present-day isotopic record. More recent modelling¹⁶ has indicated that a maximum of $3 \text{ km}^3 \text{ yr}^{-1}$ of continental crust need be subducted at present to sustain a no-growth model. This value was considered a maximum. Although crustal growth models cannot be tested by individual isotopic systems, De Paolo²⁰ has calculated minimum recycling rates necessary to reproduce Nd isotopic data. Precise models depend on how the Nd isotopic characteristics of the mantle are assumed to have evolved, but minimum present-day recycling rates of $\sim 1-1.5 \text{ km}^3 \text{ yr}^{-1}$ are indicated. The lower recycling rate of about $1 \text{ km}^3 \text{ yr}^{-1}$ assumes an exponentially decreasing $^{143}\text{Nd}/^{144}\text{Nd}$ ratio for the mantle with geological age but requires a much greater increase in recycling rate with geological age, such that the rate would have to increase by a factor of about 3 at $\sim 2,000$ Myr. The mantle evolution model giving $1.5 \text{ km}^3 \text{ yr}^{-1}$ only requires a doubling of the recycling rate at $\sim 2,000$ Myr (see ref. 20).

The cornerstone on which the no-growth, crustal recycling model has been built is that the freeboard and thickness of the continents have remained virtually constant for the past 2,500-

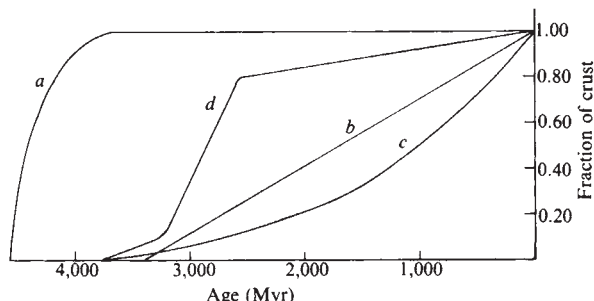


Fig. 1 Various models of continental crustal growth. *a*, Early crustal growth with subsequent recycling of continental crust through the mantle¹⁴⁻¹⁶; *b*, linear crustal growth through geological time¹; *c*, exponentially increasing crustal growth through geological time²; *d*, episodic crustal growth with the major period of crustal growth occurring between 3,200 and 2,500 Myr (ref. 13). In this model, 80–90% of the continents had formed by 2,500 Myr. *a* and *d* are consistent with the constant freeboard models of Wise^{17,18}, while *b* and *c* are not.

3,000 Myr. The constant freeboard model has gained widespread acceptance, but reasonably good records are in fact only available for the Phanerozoic^{17,18}. Also, Hallam¹⁹ has suggested that the freeboard record in fact shows a gradual continental emergence during the Phanerozoic. Such evidence does not necessarily exclude long-term constant freeboard, because continental emergence also is well documented for early and late Proterozoic times^{17-19,21}, and can be explained by continental thickening or decreasing volume of ocean ridge systems during the Phanerozoic¹⁹. In any case, if the constant freeboard model is accepted, it is clear that the continuous or accelerating crustal growth models of Hurley¹ or Hurley and Rand² are inconsistent, primarily because they fail to predict the near constant thickness of the continental crust since the late Archaean (see refs 17, 18, 22). However, most workers no longer support models of constant or accelerating crustal growth rates. Instead, models of episodic crustal growth have gained favour and conventional interpretations of the isotopic and geochemical data indicate at least 70–90% of the continental crust had formed by 2,500 Myr (refs 7–13). Recently we have used changes in trace element distributions in sedimentary rocks to model changes in area of continental crust during the late Archaean¹³. Such modelling, in conjunction with an orthodox interpretation of the isotopic record, indicates a massive increase in crustal volume at this time, such that 65–75% of the continental crust formed during the period 3,200–2,500 Myr. It was also shown that such a model was entirely consistent with constant freeboard during the past 2,500–3,000 Myr. Consequently, constant freeboard does not uniquely support models of crustal evolution involving very early growth and subsequent recycling (Fig. 1).

In no-growth models, the most interesting mechanism of recycling is the subduction of sediments of continental crustal origin¹⁴⁻¹⁶. Hydrothermal alteration of the oceanic crust may be an efficient mechanism for transport of K and U (refs 23, 24) but probably does not significantly affect other elements of isotopic significance, such as Sm, Nd, Lu, Th and Hf (see ref. 25). More difficult to assess is the process of tectonic erosion during subduction^{26,27}, but it appears to be minor²⁶. Evidence from Nd (ref. 28) and Be (ref. 29) isotopes clearly indicates that in some subduction zones, small amounts of sediment are subducted to depths sufficient for their involvement in arc volcanism. However, are sufficient amounts of the continental crust (up to 3 km³ yr⁻¹) delivered to the oceans via sedimentation in such volumes to supply the subduction rate required by recycling models?

The mass of sediment entering the ocean has been estimated in two ways, either by measuring the rate of denudation of the continents^{30,31} or by estimating the actual sediment flux in rivers, glaciers and winds³²⁻³⁴. Estimates of sediment transfer to the oceans based on denudation rates, are typically two to four times the estimates based on measured sediment flux. Sediment

transfer estimates based on denudation rates are almost certainly inaccurate because they have generally failed to take into account the large proportion of sediment which never reaches the ocean but is stored in intracontinental traps such as stream valleys and flood plains^{34,35}. On the other hand, recent estimates of river sediment loads allow reasonable estimates of sediment supply to the oceans.

The greatest proportion of material delivered to the oceans from the continents is via rivers, either as the particulate load or in solution. Milliman and Meade³⁴ have estimated the total particulate load (suspended, bedload, flood discharge) at about 16×10^{15} g yr⁻¹—a value somewhat less than many previous estimates^{32,36}. Estimates of the dissolved load in rivers (including groundwaters) are $\sim 4.5 \times 10^{15}$ g yr⁻¹ (ref. 33). Other inputs include glacial erosion, marine erosion and aeolian material, which in total account for $\sim 2.4 \times 10^{15}$ g yr⁻¹ (ref. 33). Return flux via recycling of sea salts amounts to about 0.3×10^{15} g yr⁻¹ (ref. 33). Thus the total flux of crustal material to the oceans, is 22.6×10^{15} g yr⁻¹. Assuming an average crustal density of 2.8 g cm⁻³ results in a volume flux of ~ 8.1 km³ yr⁻¹ of crustal material.

Since the beginning of agriculture, about 17,000 yr ago³⁷, man's influence has caused erosion rates to increase by as much as one to three orders of magnitude in some places^{35,38-40}. The influence of this activity on the sediment flux into the oceans, although less due to storage in river valleys, flood plains and estuaries, is still large. Douglas³⁸ suggested that man may have increased sedimentation rates by as much as a factor of 10. In a study of the Atlantic-draining rivers of eastern USA, Meade⁴¹ found that human activity increased sedimentation rates by a factor of 4–5. From a global analysis, Judson^{39,40} suggested that human influence increased overall sediment yield to the oceans by a factor of 2–2.5. For this analysis, we adopt the lowest value of 2.0 and assume it applies only to the total riverine flux and aeolian flux. The material flux from glacial and marine erosion and the recycling of sea salts is assumed to be constant. This approach is conservative and realistic for three reasons: (1) we have adopted the lowest factor of human influence on sedimentation rates and applied it only to riverine and aeolian fluxes; (2) many areas of the world containing rivers which have the largest sediment yields, such as South-East Asia, China and India, have a long history of intense cultivation; (3) our estimate agrees with other independent estimates of pre-human sediment yield rates (see below). In this case, the material flux from continents to oceans before the interference of man is unlikely to exceed 12.3×10^{15} g yr⁻¹ or 4.4 km³ yr⁻¹ of crustal material.

Judson³⁹ examined erosion rates in areas relatively unaffected by man's influence and by extrapolation, estimated a global pre-man sediment yield of 9.3×10^{15} g yr⁻¹ or 3.3 km³ yr⁻¹. Using the mass balance of dissolved Na in rivers, Gregor⁴² estimated a flux of 10.5×10^{15} g yr⁻¹ or 3.8 km³ yr⁻¹ (this value presumably excludes aeolian, glacial and marine erosion which would revise the values to 12.9×10^{15} g yr⁻¹ or 4.6 km³ yr⁻¹). Garrels *et al.*'s⁴³ model of the Phanerozoic sediment mass-age distribution pattern assumed constant sediment mass and sedimentary recycling (intracrustal recycling) rates. Such a model is consistent with an average Phanerozoic sedimentation rate of 7.5×10^{15} g yr⁻¹ or 2.7 km³ yr⁻¹. Using similar methods, Gregor⁴⁴ more recently estimated an average sedimentation rate of 10.0×10^{15} g yr⁻¹ or 3.6 km³ yr⁻¹ since the Devonian. These values are compared with our estimates in Table 1. The value of 12.3×10^{15} g yr⁻¹ sediment yield deduced in this study is, in fact, higher than most other estimates, but there is reasonable agreement.

Of the material delivered to the oceans, much never reaches the true oceanic crust, from where it may ultimately be subducted, but is deposited on continental crust. Sediment accumulation rates on modern continental terraces are >0.003 cm yr⁻¹ (ref. 45). If we assume this rate for shelves and slopes, assume that continental crust extends to the slope-rise break ($\sim 2,300$ m depth)⁴⁶ and adopt a sediment density of 2.0 g cm⁻³, then at least 3.9×10^{15} g yr⁻¹ or 1.4 km³ yr⁻¹ of crustal material must

Table 1 Estimates of pre-man annual global sediment yield to oceans

Mass (10^{15} g)	Crustal volume (km^3)*	Ref.	Method of estimate
9.3	3.3	39	Global extrapolation of erosion rates of areas relatively unaffected by man.
12.9†	4.6	42	Sodium mass balance in rivers.
10.0	3.6	40	Method not explicitly stated.
7.5‡	2.7	43	Model assuming constant sedimentary mass and sedimentary recycling rate for Phanerozoic time.
10.0§	3.6	44	Similar to ref. 43.
12.3	4.4	This study	Uses measured sedimentation rates and estimates of the influence of man.

* Assumes crustal density of 2.8 g cm^{-3} .

† Includes $2.4 \times 10^{15} \text{ g}$ to account for glacial, aeolian and marine erosion.

‡ Represents a Phanerozoic average.

§ Represents average since Devonian.

be deposited on the continental crust. Thus, no more than $3.0 \text{ km}^3 \text{ yr}^{-1}$ of crustal material is deposited on oceanic crust.

Our understanding of sedimentation in relation to plate tectonics has advanced considerably⁴⁷. The most important source of sediment to the oceans is rivers. Big rivers appear to form in specific tectonic settings which commonly are unrelated to subduction in space and time^{47,48}. Thus many major river systems form at passive continental margins (such as the Amazon, Mississippi or Niger) or adjacent to continental collision basins (such as the Ganges or Indus). In these environments, extremely large submarine sediment cones (such as the Amazon cone or Indus cone) form at the continental crust-oceanic crust interface. Little of this sediment reaches the floor of the Pacific where subduction is currently occurring. These observations could explain some of the great disparity in sediment volumes for the Atlantic and Pacific basins, which has long puzzled geologists^{24,30,49}. The result is that immense accumulations of sediment can be deposited on oceanic crust where subduction is not taking place. It seems unlikely that such thicknesses of sediment (up to 12–15 km) could ever be removed efficiently by some later episode of subduction, given the graben-fill mechanism of sediment subduction that is generally invoked^{16,27,50,51}. It is difficult to make a confident global estimate of the volumes of material accumulating in these environments, although sediment distribution maps suggest it is large. For example, the Bengal Fan alone has accreted at an average rate of almost $0.4 \text{ km}^3 \text{ yr}^{-1}$ over the past 55 Myr (ref. 47).

From this analysis, we conclude that no-growth crustal recycling models involving anything near present-day recycling rates of $3 \text{ km}^3 \text{ yr}^{-1}$ are highly unlikely because they infer quantitative removal of all sediment being deposited on oceanic crust. Given the abundance of deep marine sedimentary rocks in the geological record, such efficient removal of sediment via subduction is highly improbable.

Even if we accept that as little as $1.0\text{--}1.5 \text{ km}^3 \text{ yr}^{-1}$ of crustal material needs to be subducted to satisfy a no-growth model, as suggested by Nd-isotopic evidence²⁰, there are still severe mass balance problems. It has been suggested that the bulk of crustal material subducted is pelagic sediment^{26,52}. There is geophysical evidence that the physical properties of diagenetically altered, biogenic, pelagic sediments are similar to those of submarine-weathered oceanic basalt⁵². However, if we consider selective subduction of the biogenic carbonate and siliceous pelagics⁵², the volume required to satisfy the no-growth model increases dramatically. One argument commonly cited for subduction of pelagic sediments, on a large scale, is the lack of such deposits in ancient eugeosynclinal successions^{53,54}. This argu-

ment requires caution because pelagic sediments ultimately would be expected to be preserved at continental margins where later erosion and redeposition are most likely to occur.

Recent estimates of pelagic sedimentation rates are $<3.0 \times 10^{15} \text{ g yr}^{-1}$ (ref. 55). Such a value appears to be consistent with the findings of the GEOSECS program⁵⁶, and includes both terrigenous and biogenic carbonate and siliceous material. At average crustal density, this is equivalent to about $1.1 \text{ km}^3 \text{ yr}^{-1}$ of crust. However, the concentration of many elements, of isotopic importance, cannot be equated between pelagic sediments and the continental crust. The terrigenous fraction of pelagic sediments is highly variable but probably averages 60% (refs 55, 56). The remainder is comprised of biogenic siliceous and carbonate material. The concentrations of elements such as Th, U, Rb and the rare earth elements (REE) are much lower in biogenic than in terrigenous debris; for example, REE content in siliceous and carbonate biogenic material is typically 5–10 times lower than crustal abundances^{57,58}. Using these data, the sedimentation rate of crustal 'equivalent' material for the REE is $<0.8 \text{ km}^3 \text{ yr}^{-1}$. Thus, again the no-growth crustal recycling model (even when appealing to very low recycling rates) infers quantitative removal of pelagic sediments. The presence of pelagic sedimentary rocks in the geological record and the good evidence for accretion and 'subcretion' of much pelagic material at subduction zones²⁶ indicates that quantitative removal of pelagic sediments is unlikely.

A final point which must be addressed is how sedimentation rates have varied with time. For example, it might be argued that erosion and sedimentation rates were much higher before the development of land plants in the Devonian. Any such suggestion must be tempered by the fact that recycling rates must increase exponentially with geological age to sustain no-growth models. Gregor⁴² suggested that sedimentary survival rates indicated pre-Carboniferous erosion rates were four times as great as Recent (pre-agriculture) rates. However, Garrels and Mackenzie⁵⁹ pointed out that expected differences in sedimentary recycling rates explained the data equally well, and suggested Palaeozoic depositional rates may even have been less than Mesozoic-Cenozoic rates. Later modelling⁴³ proposed a Phanerozoic average sediment yield of only $7.5 \times 10^{15} \text{ g yr}^{-1}$ or less than two-thirds of the Recent pre-agriculture rates calculated in this study. Using similar models, Gregor⁴⁴ has more recently estimated a mean sedimentation rate of $10.0 \times 10^{15} \text{ g yr}^{-1}$ since the Devonian (see above and Table 1). Schopf⁴⁶ concluded that available evidence did not support any drastic changes in denudation rates.

Constant freeboard of the continents during the past 2,500–3,000 Myr is consistent with current versions of both growth and no-growth models for the continental crust. Before man's influence, the removal of material from the continental crust to oceanic crust via sedimentation probably did not exceed $3.0 \text{ km}^3 \text{ yr}^{-1}$. Accumulation rates of pelagic sediments are such that $<1 \text{ km}^3 \text{ yr}^{-1}$ of crustal equivalent for REE, Th, U, Rb is deposited. Such sedimentation rates appear to be insufficient to sustain no-growth crustal recycling models. There is little evidence for exponentially increasing sedimentation rates with geological age. Consequently, we conclude that the continental crust has grown throughout geological time but that growth was nearly complete by ~2,500 Myr (see refs 11, 13).

We thank Richard Arculus and Shen-su Sun for reading an earlier version of this paper and Gail Stewart for preparing the manuscript.

Received 6 May, accepted 20 September 1983.

- Hurley, P. M. *Geochim. cosmochim. Acta* **32**, 273–283 (1968).
- Hurley, P. M. & Rand, J. R. *Science* **164**, 1229–1242 (1969).
- Moorbath, S. in *The Early History of the Earth* (ed. Windley, B. F.) 351–360 (Wiley, New York, 1976).
- Moorbath, S. *Phil. Trans. R. Soc. A* **288**, 401–413 (1978).
- Veizer, J. in *The Early History of the Earth* (ed. Windley, B. F.) 569–578 (Wiley, New York, 1976).
- Veizer, J. & Compston, W. *Geochim. cosmochim. Acta* **40**, 905–914 (1976).
- Veizer, J. & Jansen, S. L. *J. Geol.* **87**, 341–370 (1979).
- Veizer, J., Compston, W., Hoefs, J. & Neilson, J. *Naturwissenschaften* **69**, 173–180 (1982).
- McCulloch, M. T. & Wasserburg, G. J. *Science* **200**, 1003–1011 (1978).
- Allegre, C. J. *Tectonophysics* **81**, 109–132 (1982).

11. Taylor, S. R. & McLennan, S. M. *Phil. Trans. R. Soc. A* **301**, 381–399 (1981).
12. Dewey, J. F. & Windley, B. F. *Phil. Trans. R. Soc. A* **301**, 188–206 (1981).
13. McLennan, S. M. & Taylor, S. R. *J. Geol.* **90**, 347–361 (1982).
14. Armstrong, R. L. *Rev. Geophys.* **6**, 175–199 (1968).
15. Armstrong, R. L. & Hein, S. M. *Geochim. cosmochim. Acta* **37**, 1–18 (1973).
16. Armstrong, R. L. *Phil. Trans. R. Soc. A* **301**, 443–472 (1981).
17. Wise, D. U. *Geol. Soc. Am. Mem.* **132**, 87–100 (1972).
18. Wise, D. U. in *The Geology of Continental Margins* (eds Burke, C. A. & Drake, C. L.) 45–58 (Springer, Berlin, 1974).
19. Hallam, A. *Nature* **269**, 769–772 (1977).
20. De Paolo, D. J. *Geochim. cosmochim. Acta* **45**, 1253–1254 (1981).
21. Rutland, R. W. R. *J. Proc. R. Soc. N.S.W.* **115**, 33–60 (1982).
22. Condie, K. C. *Bull. geol. Soc. Am.* **84**, 2981–2992 (1973).
23. Fyfe, W. S. *Geosci. Can.* **3**, 82–83 (1976).
24. Fyfe, W. S. *Phil. Trans. R. Soc. A* **291**, 433–445 (1979).
25. Staudigel, H. & Hart, S. R. *Geochim. cosmochim. Acta* **47**, 337–350 (1983).
26. Karig, D. E. & Kay, R. W. *Phil. Trans. R. Soc. A* **301**, 233–251 (1981).
27. Uyeda, S. *Episodes* **1983** (2), 19–24 (1983).
28. Whitford, D. J., White, W. M., Jazek, P. A. & Nicholls, I. A. *Yb Carnegie Instn., Wash.* **78**, 304–308 (1979).
29. Brown, L., Klein, J., Middleton, R., Sacks, I. S. & Tera, F. *Nature* **299**, 718–720 (1982).
30. Gilluly, J. *Spec. Pap. geol. Soc. Am.* **62**, 7–18 (1955).
31. Schumm, S. A. *U.S. geol. Surv. prof. Pap.* 545–H (1963).
32. Holeman, J. N. *Wat. Resour. Res.* **4**, 737–747 (1968).
33. Garrels, R. M. & Mackenzie, F. T. *Evolution of Sedimentary Rocks* (Norton, New York, 1971).
34. Milliman, J. D. & Meade, R. H. *J. Geol.* **91**, 1–21 (1983).
35. Meade, R. H. *J. Geol.* **90**, 235–252 (1982).
36. Kuenen, P. J. *Marine Geology* (Wiley, New York, 1950).
37. Thomas, H. A *History of the World* (Harper & Row, New York, 1979).
38. Douglas, I. *Nature* **215**, 925–928 (1967).
39. Judson, S. *Am. Scient.* **56**, 356–374 (1968).
40. Judson, S. *Geol. Soc. Am. Abstr.* **3**, 615 (1971).
41. Meade, R. H. *Bull. geol. Soc. Am.* **80**, 1265–1274 (1969).
42. Gregor, B. *Nature* **228**, 273–275 (1970).
43. Garrels, R. M., Lerman, A. & Mackenzie, F. T. *Am. Scient.* **64**, 306–315 (1976).
44. Gregor, C. A. *EOS* **64**, 344 (1983).
45. Schwab, F. L. *Geology* **4**, 723–727 (1976).
46. Schopf, T. J. M. *Paleoceanography* (Harvard University Press, Cambridge, 1980).
47. Reading, H. G. *Proc. geol. Ass.* **93**, 321–350 (1982).
48. Potter, P. E. *J. Geol.* **86**, 13–33 (1978).
49. Gilluly, J. *Bull. geol. Soc. Am.* **82**, 2382–2396 (1971).
50. Ludwig, W. J. et al. *J. geophys. Res.* **71**, 2121–2137 (1966).
51. Schweller, W. J. & Kulm, L. D. *Mar. Geol.* **28**, 271–291 (1978).
52. Moore, J. C. *Geology* **3**, 530–532 (1975).
53. Scholl, D. W. & Marlow, M. S. *Soc. Econ. Palaeont. Miner. Spec. Publ.* **19**, 193–211 (1974).
54. Scholl, D. W. & Vallier, T. L. in *Expanding Earth Symposium*, Sydney (ed. Carey, S. W.) 235–245 (University of Tasmania, Hobart, 1983).
55. Lisitsyn, A. P., Lukashin, V. N., Gurvich, Ye. G., Gordeyev, V. V. & Demina, L. L. *Geochim. Inter.* **19**, 102–110 (1982).
56. Beoecker, W. A. & Peng, T.-H. *Tracers in the Sea* (Lamont-Doherty Geological Observatory, Palisades, 1982).
57. Piper, D. Z. *Geochim. cosmochim. Acta* **38**, 1007–1022 (1974).
58. Shimizu, H. & Masuda, A. *Nature* **266**, 346–348 (1977).
59. Garrels, R. M. & Mackenzie, F. T. *Nature* **231**, 382–383 (1971).

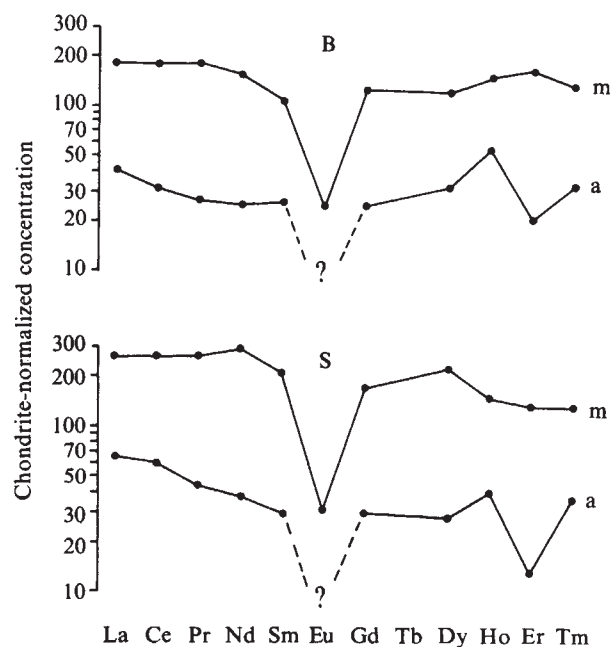


Fig. 1 REE abundances determined by ion-probe analysis in merrillite (m) and apatite (a) in the Barbotan (B) and Salles (S) chondrites, normalized with respect to mean chondritic abundances¹¹. (Yb, Tb, Lu not determined; Eu not determined in apatite.)

Rare earth elements in chondritic phosphates—implications for ²⁴⁴Pu chronology

S. J. B. Reed*, D. G. W. Smith† & J. V. P. Long*

* Department of Earth Sciences, University of Cambridge, Cambridge CB3 0EZ, UK

† Department of Geology, University of Alberta, Edmonton, Alberta, Canada T6G 2E3

In stony meteorites of the chondrite class, phosphates (comprising about 0.6% of the whole) act as host for rare-earth elements (REE) and for Pu which was present at the time of formation. For the purpose of Pu–Xe dating it has been suggested that, in the absence of a stable Pu isotope, measured Xe abundances may be related to REE concentrations, on the grounds of presumed similarity in the geochemical behaviour of Pu and REE¹. There are two phosphate phases in chondrites, apatite and merrillite (or whitlockite), which are difficult to separate from each other for analysis. Recent results based on grain-size discrimination² indicate similar REE concentrations in both phosphates, whereas Pu favours merrillite³. Direct *in situ* analyses of individual grains by the ion-probe technique, however, show REE concentrations to be about five times higher in merrillite. The use of REE in the role of Pu analogue thus appears a better prospect than suggested by the previous indirect evidence.

Most determinations of REE in chondritic phosphates^{4,5} do not discriminate between apatite and merrillite, owing to

difficulty in separation. However, Ebihara and Honda² attempted to utilize grain-size differences for this purpose (other physical properties being closely similar). Phosphate separates were divided into three size-fractions, which were analysed by instrumental neutron activation (INAA). The apatite/merrillite ratio in each fraction, as estimated from the major element concentrations, was found to be a function of grain size, enabling REE concentrations in each phosphate phase to be estimated. Little difference between the two was found in the chondrites Bruderheim (type L6) and Richardton (H5).

The problem of physical separation can be avoided by using a probe technique. The electron microprobe is insufficiently sensitive for present purposes, but ion-probe analysis⁶ holds out the prospect of much lower detection limits, though with somewhat inferior spatial resolution. The ion-probe measurements reported here were carried out with an extensively modified AEI IM-20 instrument^{7,8}, using a 20-μm beam of ¹⁶O⁺ ions to bombard the specimen (a carbon-coated polished thin section). Previously identified phosphate grains were located with the aid of the built-in microscope and the relevant peaks in the mass spectrum were recorded using the computer-controlled mass spectrometer.

While maximum sensitivity is obtained with low mass resolution, when the transmission of the mass spectrometer is greatest, the REE peaks are then subject to interference from isobaric molecular ions. Light REE monoxides interfere with the atomic peaks of heavy REE (from Gd upwards), but this difficulty can be avoided by using heavy REE monoxide peaks for analytical purposes⁹. However, interferences from molecules composed of matrix atoms (Ca, P and O for example) can occur at any mass and their intensity is difficult to predict. Mass defect calculations indicate that a mass resolution of 6,000 (defined as mass divided by peak width at the 10% level, in mass units) is sufficient to resolve nearly all likely matrix interferences, but the reduction in transmission at such resolution rules out this approach for peaks of low intensity.

The strategy used here for dealing with interferences was as follows. In the case of poly-isotopic REE, the intensities of the major isotopes, measured at low resolution (about 400), were compared with the known relative abundances, and the isotope showing the least interference was selected. The chosen isotopes,

Table 1 Major-element concentrations (wt %) in phosphates

	Durango fluorapatite	Typical chondritic chlorapatite	merrillite
O	38.3	38.9	41.1
F	3.5	0.4	<0.1
Na	0.2	0.3	2.0
Mg	<0.1	0.5	2.4
P	17.8	17.2	19.6
Cl	0.4	4.4	<0.1
Ca	38.6	36.1	32.6
Fe	<0.1	0.8	1.0

together with those of the mono-isotopic REE, were inspected at a mass resolution of 6,000, intensity permitting. Interferences were found to be negligible (<10%) for light REE (La–Sm) in merrillite, but were more significant (up to 70% of the peak intensity) in apatite. Where necessary, interference correction factors derived from the high-resolution spectra were applied to the low-resolution data. Amongst the heavy REE, serious interferences were observed in the case of Yb, Tb and Lu, which therefore were omitted from the analysis. Low intensity prevented interference correction factors of reasonable accuracy being determined for the other heavy REE, results for which are hence given in uncorrected form. Data for Eu in merrillite were obtained, but not in apatite, where the concentration was too low.

REE peak intensities measured at low resolution, corrected for interference where appropriate and normalized relative to ^{44}Ca , were compared with measurements on a fluorapatite previously analysed by INAA. REE concentrations in the chondritic phosphates were calculated on the assumption that matrix effects could be neglected. The compositions of the phases concerned are given in Table 1. The difference between the chondritic chlorapatite and the fluorapatite standard with respect to relative Cl and F content is most unlikely to have a significant effect on REE ion yields. There is a greater difference between fluorapatite and merrillite, but this is still unlikely to have more than a minor effect: thus the conclusions discussed below cannot be seriously disputed on the grounds of possible matrix effects.

Chondrite-normalized REE concentrations determined with the ion probe for apatite and merrillite in Barbotan (H5 type) and Salles (H6) are shown in Fig. 1. The merrillite patterns are very similar to those obtained previously^{4,5} by bulk analysis of separated phosphate (probably mostly merrillite), being essentially flat, apart from the negative Eu anomaly (assumed to be due to divalent Eu entering feldspar). The apatite pattern is also more or less flat (allowing for the likely effect of uncorrected interferences on the heavy REE), but the absolute REE concentrations in apatite are lower by a factor of about five.

These results are in marked disagreement with those obtained by the indirect size-fraction method²: the reason is uncertain, but errors of this magnitude in the ion-probe data are implausible, in spite of the admitted limitations of this relatively new technique. Note that whitlockite in lunar basalt shows a similar enrichment in REE compared with apatite, the concentrations in this case being within the range of electron microprobe analysis¹⁰. Evidently the possibility of using REE in chondritic Pu–Xe studies should not be dismissed, since the partitioning of REE between the two phosphate phases appears similar to that of Pu.

Support for the Cambridge Ion Probe Laboratory from the Natural Environment Research Council is gratefully acknowledged; we also thank Dr P. Potts (Open University) for the INAA analysis of the Durango apatite standard.

Received 28 July; accepted 14 September 1983.

1. Lugmair, G. W. & Marti, K. *Earth planet. Sci. Lett.* **35**, 273–284 (1977).
2. Ebihara, M. & Honda, M. *Earth planet. Sci. Lett.* **63**, 433–445 (1983).
3. Crozaz, G. *Earth planet. Sci. Lett.* **23**, 164–169 (1974).
4. Mason, B. & Graham, A. L. *Smithson. Contr. Earth Sci.* **3**, 1–17 (1970).
5. Allen, R. O. & Mason, B. *Geochim. cosmochim. Acta* **37**, 1435–1456 (1973).
6. Shimizu, N. & Hart, S. R. A. *Rev. Earth planet. Sci.* **10**, 483–526 (1982).

7. Banner, A. E. & Stimpson, B. P. *Vacuum* **24**, 511–517 (1974).
8. Long, J. V. P. *et al.* in *X-ray Optics and Microanalysis* (eds Beaman, D. R. *et al.*) 316–321 (Pendell, Midland, Michigan, 1980).
9. Reed, S. J. B. in *Microbeam Analysis—1981* (ed. Geiss, R. H.) 87–90 (San Francisco Press, 1981).
10. Griffin, W. L. *et al.* *Earth planet. Sci. Lett.* **15**, 53–58 (1972).
11. Evensen, N. M. *et al.* *Geochim. cosmochim. Acta* **42**, 1199–1212 (1978).

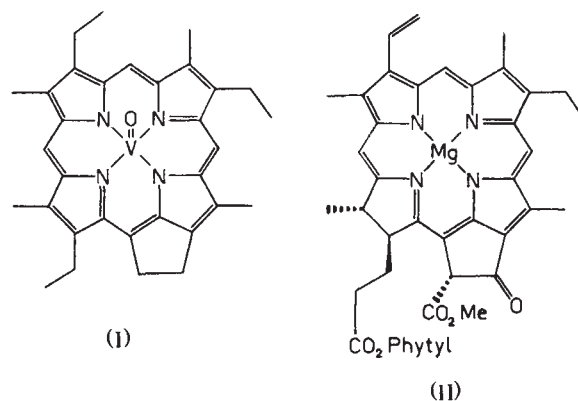
Determination of the crystal structure of a petroporphyrin isolated from oil shale

A. Ekstrom, C. J. R. Fookes, T. Hambley, H. J. Loeh, S. A. Miller & J. C. Taylor

CSIRO, Division of Energy Chemistry, Lucas Heights Research Laboratories, Private Mail Bag 7, Sutherland, New South Wales 2232, Australia

The isolation and identification of the metal porphyrins found in crude oils and oil shales have been extensively studied since their discovery in these materials by Treibs¹. He postulated that the principal compound found in geological samples is vanadyl deoxophylloerythroetioporphyrin (DPEP) and developed a detailed scheme² to show that this compound (I) was formed from chlorophyll (II). Consequently, the presence of metal porphyrins in crude oils was taken to show their biological origin, and his work thus laid the foundations for organic geochemistry. However, despite its general acceptance, proof of Treibs' hypothesis still requires the isolation and unambiguous structure determination of vanadyl-DPEP from geological sources. Thus, although C₃₂ DPEP has been synthesized on several occasions^{3–6}, its identity with the naturally occurring material has not been established. We now report that the determination of the crystal structure of a petroporphyrin isolated from an Australian oil shale confirms the above hypothesis.

Mass spectrometric examinations of petroporphyrins have consistently shown them to be very complex mixtures^{7–9}, and the assignment of the structure of these compounds to DPEP derivatives has been based solely on visible spectroscopy and



molecular weight data. To our knowledge, the only exception to this approach has been the recent work of Quirke *et al.*¹⁰, who examined a demetallated C₃₂ DPEP compound from gilsonite by NMR techniques. The results were consistent with three possible DPEP isomers, but no conclusive identification was possible. Their studies involved the demetallation of the metal porphyrins in very harsh acid conditions. However, it is known¹¹ that substantial decomposition of the porphyrins can occur during demetallation and our experience has confirmed this observation. The occurrence of this decomposition clearly raises the question of whether the isolated metal-free porphyrins accurately reflect the structure of the metal porphyrins originally present in the sample. We now report the first unequivocal

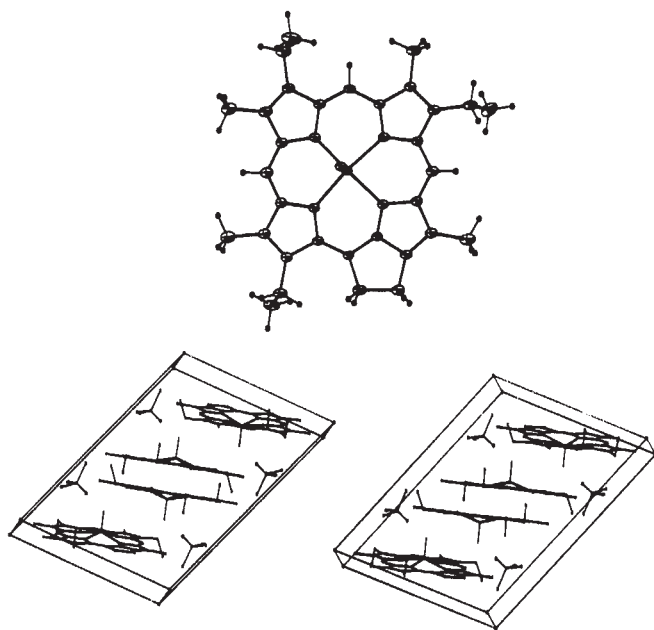


Fig. 1 Structure of the vanadyl-DPEP isolated from oil shale.

identification by a complete X-ray determination of a vanadyl porphyrin isolated from a geological sample.

Crushed oil shale from the Julia Creek Deposit in North Queensland, Australia, was extracted with chloroform at room temperature. The crude extract was chromatographed on Merck Kieselgel 60 using a gradient of CHCl_3 in CCl_4 . The band containing the vanadium porphyrins was collected, and the porphyrins precipitated by the addition of cold pentane. The HPLC (C18 column, CH_3OH solvent) of this product showed seven major peaks. The predominant peak was isolated on a semipreparative column and mass spectrometry of this fraction indicated two predominant molecular weights at 541 and 527 AMU. Preparative TLC (silica, C_6H_6) of the above fraction gave two bands of which the less polar (molecular weight 541 AMU) was readily obtained as single crystals from $\text{CHCl}_3/\text{CH}_3\text{OH}$ solution.

The crystal chosen for X-ray analysis was monoclinic, space group $\text{P2}_1/\text{c}$ with $a = 12.912 \pm (3)$, $b = 14.151 \pm (4)$ and $c = 18.404 \pm (8)$, $\beta = 70.34 \pm (2)$. The structure was refined by block-matrix least-squares techniques to an R factor (on $3,220 F$, $I > 2\sigma(I)$ four-circle diffractometer data) of 0.077, with $R = \Sigma(|F_o| - |F_c|)/\Sigma|F_o|$. The structure obtained is shown in Fig. 1 and is seen to be the vanadyl C_{32} DPEP structure originally proposed by Treibs¹, and shown schematically as compound (I). A full description of the X-ray diffraction results will be given elsewhere.

This result is important because it represents the first complete identification of a petroporphyrin unlikely to have been chemically altered by the extraction procedure. More importantly, however, the determination of this structure confirms, for the first time, the long-held belief that petroporphyrins are derived from chlorophyll.

We thank Mr A. Muradian of CSR Ltd for the sample of oil shale used in this study.

Received 29 July; accepted 25 August 1983.

1. Treibs, A. *Ann. Chem.* **510**, 42–62 (1934).
2. Treibs, A. *Angew. Chem.* **49**, 682–686 (1936).
3. Fischer, H. & Hofman, H. *J. Ann. Chem.* **517**, 274–277 (1935).
4. Sugihara, J. M. & McGee, L. R. *M. J. org. Chem.* **22**, 795–798 (1957).
5. Baker, E. W., Corwin, A. H., Klesper, E. & Wei, P. E. *J. org. Chem.* **33**, 3144–3148 (1968).
6. Chaudry, I. A., Clezy, P. S. & Mizra, A. H. *Aust. J. Chem.* **33**, 1095–1104 (1980).
7. Baker, E. W. *J. Am. chem. Soc.* **88**, 2311–2315 (1966).
8. Baker, E. W., Yen, T. F., Dickie, J. P., Rhodes, R. E. & Clark, L. F. *J. Am. chem. Soc.* **89**, 3631–3639 (1967).
9. Thomas, D. W. & Blumer, M. *Geochim. cosmochim. Acta* **28**, 1147–1154 (1964).
10. Quirke, J. M. E., Maxwell, J. K., Eglington, G. & Sanders, J. K. M. *Tetrahedron Lett.* **21**, 2987–2990 (1980).
11. Baker, E. W. in *Organic Geochemistry, Methods and Results* (eds Eglington, G. & Murphy, M. T. J.) 464–497 (Springer, Berlin, 1969).

Endothelium-dependent relaxation in rat aorta may be mediated through cyclic GMP-dependent protein phosphorylation

Robert M. Rapoport, Martin B. Draznin & Ferid Murad

Departments of Medicine and Pharmacology, Stanford University School of Medicine, Palo Alto Veterans Administration Medical Center, 3801 Miranda Avenue, Palo Alto, California 94304, USA

The action of some vascular smooth muscle relaxants depends on the presence of the endothelium^{1–10}. We have recently shown that relaxation may be mediated through the formation of cyclic GMP¹⁰. The nitrovasodilators are another class of relaxants which exert their effects through the formation of cyclic GMP^{11–13}, although their relaxation is independent of the presence of the endothelium^{1,2,10}. Their relaxant properties seem to depend on free radical formation—specifically, the formation of nitric oxide^{14,15}. The NO-induced smooth muscle relaxation is proposed to occur through activation of guanylate cyclase and the formation of cyclic GMP^{11–13}. Protein phosphorylation is thought to be a common event in the pathway for many biological phenomena¹⁶. Moreover, sodium nitroprusside and 8-bromo cyclic GMP induce similar patterns of protein phosphorylation in intact rat thoracic aorta¹⁷. Here we report that the patterns of protein phosphorylation induced by the endothelium-dependent vasodilators and nitrovasodilators were identical. Incorporation of ³²P into myosin light chain was decreased by both classes of agents. Removal of the endothelium abolished the changes in phosphorylation with the endothelium-dependent vasodilator (acetylcholine), but not those with the nitrovasodilator (sodium nitroprusside). These results suggest that endothelium-dependent vasodilators and nitrovasodilators induce relaxation through cyclic GMP-dependent protein phosphorylation and dephosphorylation of myosin light chain.

Sodium nitroprusside (50 nM) and acetylcholine (10 μM) caused identical patterns of protein phosphorylation in rat thoracic aorta with intact endothelium (Fig. 1). Calcium ionophore A23187 (0.3 μM) also induced an identical pattern of phosphorylation to nitroprusside and acetylcholine (data not shown). These changes occurred within three regions of the autoradiographs and are designated as groups 1, 2 and 3 (see Fig. 1 of ref. 17 for localization of these groups within the autoradiographs). The alterations in ³²P-incorporation in the proteins of groups 1 and 3 were observed in both soluble and particulate fractions, even when the particulate fraction was washed twice with homogenization buffer. The changes observed with group 2 proteins were only observed in the soluble fraction.

The changes induced by nitroprusside (50 nM) in the incorporation of ³²P into proteins in the three groups were of about the same magnitude as those induced by 10 μM acetylcholine (Fig. 1). This is consistent with the observation that after contraction with 0.3 μM noradrenaline, 50 nM nitroprusside and 10 μM acetylcholine, present for 2 min, induced relaxations of 90 and 80%, respectively. The pattern of protein phosphorylation with 10 μM acetylcholine was similar at 30 s and 2 min. Table 1 lists the molecular weights (M_r s) and pI values of proteins whose ³²P incorporation was altered.

The pattern of protein phosphorylation in groups 1 and 2 with 10 μM acetylcholine was the same in noncontracted tissues as in tissues contracted with 0.3 μM noradrenaline. Proteins 3a and 3b were identified as myosin light chain using partially purified protein standards and transfer of proteins from gels to nitrocellulose paper and recognition with antibody (Western

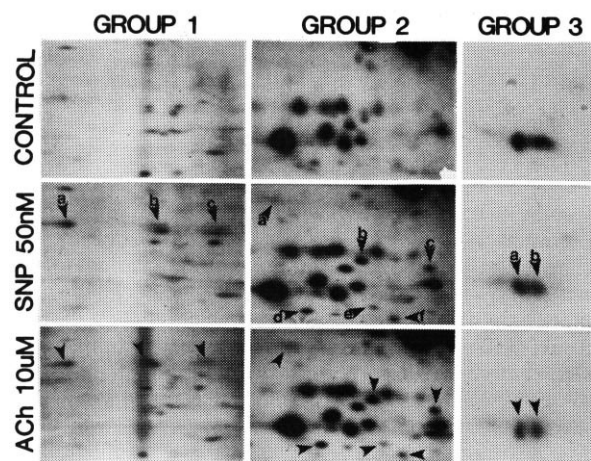


Fig. 1 Effect of sodium nitroprusside and acetylcholine on ^{32}P incorporation. Tissues were incubated with $0.3\ \mu\text{M}$ noradrenaline for 1 min followed by either $50\ \text{nM}$ sodium nitroprusside (SNP) or $10\ \mu\text{M}$ acetylcholine (ACh) for 2 min. Additional tissues remained exposed to noradrenaline for 3 min. Two-dimensional gel electrophoresis and autoradiography were then performed. 'Group' represents an area of the autoradiograph as designated in Fig. 1 of ref. 17. Molecular weights and pI values of the individual proteins are listed in Table 1.

Methods: Thoracic aortas from decapitated male, 250–300 g Sprague-Dawley rats were placed in Krebs-Ringer bicarbonate solution containing 10 mM glucose and gassed with 95% O_2 –5% CO_2 . Spiral strips were prepared as described elsewhere¹⁹. The intimal surface of some of the strips was gently rubbed with a scalpel, which removed the endothelium and left the internal elastic lamina intact as confirmed by scanning electron microscopy¹⁰. Strips derived from 2.6 aortas were cut into 1.5-cm segments and incubated in 1.5 ml of Krebs-Ringer bicarbonate with 10 mM glucose and 2 mCi of ^{32}P -orthophosphoric acid as described previously¹⁷. Agents were added for the appropriate time, after which the tissues were removed from the incubation and frozen in liquid nitrogen. Strips were homogenized in buffer containing 100 mM NaF, 80 mM sucrose, 10 mM EDTA and 10 mM *N*-tris(hydroxymethyl) methyl-2-aminoethane sulphonic acid (pH 7.4). The inclusion of $30\ \mu\text{M}$ phenylmethylsulphonyl fluoride, 1 mM sodium bisulphite, 1 mM phenanthroline, 1 mM benzamide, and $50\ \mu\text{g ml}^{-1}$ of each of soybean trypsin inhibitor, leupeptin, aprotinin and ovomucoid trypsin inhibitor during homogenization and processing did not alter the protein or autoradiographic profile of two-dimensional gels (data not shown). Homogenates were centrifuged at $105,000g$ for 60 min, the supernatant fractions removed and the pellets resuspended in fresh homogenization buffer. Before electrophoresis, aliquots of the samples were assayed for ^{32}P and protein concentration²⁰. Some buffer was added so that the ^{32}P and protein concentrations were equivalent in all supernatant or particulate fractions. Two-dimensional gel electrophoresis and autoradiography were performed as described elsewhere^{17,21}; the second dimension used 12% polyacrylamide slab gels. The use of 7% gels increased the resolution in the higher molecular weight range; however, no additional phosphorylated proteins were observed (data not shown). Typically, four tissue incubations (control and three experimental conditions) were performed for each experiment. Experiments with acetylcholine on protein phosphorylation in tissues without and with endothelium were performed four and nine times, respectively. The photographs presented are representative of the changes observed and are from paired experiments. Changes in phosphorylation were determined visually.

blot technique¹⁸). Phosphorylation of proteins 3a and 3b is increased on exposure to $0.3\ \mu\text{M}$ noradrenaline¹⁸, and a decrease in phosphorylation with acetylcholine is more apparent when the level of phosphorylation is increased by noradrenaline. Thus, these results suggest that endothelium-dependent vasodilators may induce relaxation through dephosphorylation of myosin light chain, as appears to be the case with other vasodilators^{17,18}.

Table 1 Molecular weights and pI values of proteins affected by acetylcholine and sodium nitroprusside

Proteins with increased ^{32}P	$M_r \times 10^{-3}$	pI
1a	49.0	7.8
1b	49.0	7.4
1c	49.0	7.1
2a	28.0	6.8
2b	24.0	6.5
2c	24.0	6.2
2d	21.5	6.6
2e	21.5	6.4
2f	21.3	6.4
Proteins with decreased ^{32}P		
3a	22.0	5.3
3b	22.0	5.1

Acetylcholine ($10\ \mu\text{M}$) and sodium nitroprusside ($50\ \text{nM}$) altered ^{32}P incorporation into the proteins listed. The number and letter in 'Protein' column refer to the location in Fig. 1.

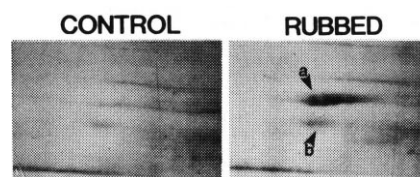


Fig. 2 Protein phosphorylation in rat thoracic aorta with and without endothelium. The endothelium from rat thoracic aorta was removed and the tissues incubated in ^{32}P and exposed to $0.3\ \mu\text{M}$ noradrenaline for 1 min and $10\ \mu\text{M}$ acetylcholine for 2 min. Two-dimensional gel electrophoresis and autoradiography were then performed. The soluble fraction is shown. See text for discussion of changes.

Removal of the endothelium abolished the changes in phosphorylation induced by acetylcholine and resulted in a protein phosphorylation pattern indistinguishable from that of the controls shown in Fig. 1. In contrast, the sodium nitroprusside-induced pattern of protein phosphorylation was unaltered by removal of the endothelium (Fig. 1 and ref. 17). This is consistent with the observation that removal of the endothelium abolished acetylcholine-induced relaxation and the associated increased levels of cyclic GMP, while relaxation and increased levels of cyclic GMP with nitroprusside were independent of the presence of the endothelium¹⁰.

Removal of the endothelium caused the phosphorylation of a soluble protein of $120,000\ M_r$ and midpoint $pI \sim 6.7$ (a, Fig. 2). However, acetylcholine ($10\ \mu\text{M}$) and/or noradrenaline ($0.3\ \mu\text{M}$) had no effect on the phosphorylation of this protein in aortic strips from which the endothelium had been removed. The mechanism by which the protein is phosphorylated and its identity are unknown. Except for protein a in Fig. 2, the protein phosphorylation profiles of aortic strips with and without endothelium were, within the current limit of resolution, identical.

Acetylcholine, sodium nitroprusside, 8-bromocyclic GMP and dibutyryl cyclic AMP also increased phosphorylation of a protein of $95,000\ M_r$ and $pI \sim 6.7$ in aortic strips with or without endothelium (b, Fig. 2; some data not shown). Although this change was observed in many experiments, for unknown reasons this was not a consistent observation. The change was usually observed in both soluble and particulate fractions.

The present study supports and extends the hypothesis that endothelium-dependent vasodilation may be mediated through the formation of cyclic GMP. Identical patterns of protein phosphorylation were induced by acetylcholine, calcium ionophore A23187 and nitroprusside, suggesting that these agents elevate the same functional pool of cyclic GMP. Furthermore, these results suggest that nitroprusside and other nitrovasodilators induce relaxation through bypassing part of

the pathway used by endothelium-dependent vasodilators and directly activate smooth muscle guanylate cyclase. Relaxation induced by the endothelium-dependent vasodilators and nitrovasodilators may be mediated through cyclic GMP-dependent dephosphorylation of myosin light chain.

We thank G. A. Martinez for technical assistance and K. A. Chewey for preparation of the manuscript. This work was supported by NIH grants AM 30787 and HL 28474, grants from the Veterans Administration and the Council for Tobacco Research—USA, Inc. and National Research Service Award GM 07673 (to R.M.R.).

Received 7 June; accepted 23 August 1983.

1. Furchgott, R. F. & Zawadzki, J. V. *Nature* **288**, 273–276 (1980).
2. Furchgott, R. F. *Trends pharmac. Sci.* **2**, 173–176 (1981).
3. Altura, B. M. & Chand, N. *Br. J. Pharmac.* **74**, 10–11 (1981).
4. Chand, N. & Altura, B. M. *Science* **213**, 1376–1379 (1981).
5. De Mey, J. G. & Vanhoutte, P. M. *J. Physiol., Lond.* **316**, 347–355 (1981).
6. Cherry, P. D., Furchgott, R. F., Zawadzki, J. V. & Jothianandan, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2106–2110 (1982).
7. De Mey, J. G. & Vanhoutte, P. M. *Circulation Res.* **51**, 439–447 (1982).
8. De Mey, J. G., Claeys, M. & Vanhoutte, P. M. *J. Pharmac. exp. Ther.* **222**, 166–173 (1982).
9. Ku, D. D. *Science* **218**, 576–578 (1982).
10. Rapoport, R. M. & Murad, F. *Circulation Res.* **52**, 352–357 (1983).
11. Katsuki, S., Arnold, W., Mittal, C. & Murad, F. *J. Cyclic Nucleotide Res.* **3**, 23–25 (1977).
12. Katsuki, S. & Murad, F. *Molec. Pharmac.* **13**, 330–341 (1977).
13. Schultz, K.-D., Schultz, K. & Schultz, G. *Nature* **256**, 750–751 (1977).
14. Arnold, W. P., Mittal, C. K., Katsuki, S. & Murad, F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3203–3207 (1977).
15. Katsuki, S., Arnold, W., Mittal, C. & Murad, F. *J. Cyclic Nucleotide Res.* **3**, 23–25 (1977).
16. Greengard, P. *Science* **199**, 146–152 (1978).
17. Rapoport, R. M., Draznin, M. B. & Murad, F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6470–6474 (1982).
18. Draznin, M. B., Rapoport, R. M., Martinez, G. A. & Murad, F. *Clin. Res.* **31**, 466A (1983).
19. Furchgott, R. F. & Bhadrakom, S. *J. Pharmac. exp. Ther.* **108**, 129–143 (1953).
20. Lowry, O. H., Rosebrough, A. L., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
21. Anderson, N. G. & Anderson, N. L. *Analyt. Biochem.* **85**, 331–340 (1978).

Anatomical distributions of four pharmacologically distinct ³H-L-glutamate binding sites

Daniel T. Monaghan, Vicky R. Holets, Darryl W. Toy & Carl W. Cotman*

Department of Psychobiology, University of California, Irvine, California 92717, USA

Glutamate is thought to serve as a major excitatory neurotransmitter throughout the central nervous system (CNS)^{1,2}; electrophysiological studies indicate that its action is mediated by multiple receptors. Four receptors have been characterized by their selective sensitivity to *N*-methyl-D-aspartate (NMDA), kainic acid (KA), quisqualic acid (QA) and 2-amino-4-phosphonobutyric acid (APB)^{1,3–5}. Electrophysiological evidence indicates that these receptors are all present in the rat hippocampus and that the anatomically discrete synaptic fields within the hippocampus exhibit differential sensitivity to the selective excitatory amino acid agents^{3,6,7}. Thus, we have used the hippocampus as a model system to investigate possible subpopulations of ³H-L-glutamate binding sites. By using quantitative autoradiography, the pharmacological specificity of ³H-L-glutamate binding in discrete terminal fields was determined. We report here that there are at least four distinct classes of ³H-L-glutamate binding sites which differ in their anatomical distribution, pharmacological profile and regulation by ions. Two of these sites seem to correspond to the KA and NMDA receptor classes, and a third site may represent the QA receptor. The fourth binding site does not conform to present receptor classifications. None of these binding sites corresponds to the major glutamate binding site observed in biochemical studies^{8–12}.

* To whom reprint requests should be addressed.

Autoradiograms prepared in the absence of Ca²⁺, Cl[−] and Na⁺ ions (Fig. 1a) reveal three populations of ³H-L-glutamate binding sites, the major population being readily displaced by low concentrations of NMDA (Fig. 1b). The distribution of these sites is most easily seen in Fig. 1c. These sites are found predominantly in the stratum oriens and stratum radiatum of the CA1 field and the inner portions of the dentate gyrus molecular layer. Moderate levels are found in the stratum oriens and stratum radiatum of CA3, the outer molecular layer of the dentate gyrus and the stratum lucidum and the dentate gyrus hilus (quantitative values are given in Fig. 1 legend). NMDA exhibits high affinity for this class of binding site. Within the stratum radiatum an NMDA inhibition constant of $1.55 \pm 0.36 \mu\text{M}$ ($\pm \text{s.e.m.}$, $n = 3$) was obtained. The potent NMDA antagonist D-2-amino-5-phosphonovaleric acid (D-APV) has an affinity constant (K_i) of $2.3 \pm 0.2 \mu\text{M}$ ($n = 2$), which corresponds well with its physiological potency (apparent $K_d = 1.4 \mu\text{M}$)⁴. Compounds which interact with the NMDA receptor¹ (ibotenate, D- α -amino adipate, and D- and L-aspartate) are potent displacers of the NMDA-sensitive ³H-L-glutamate binding (50% inhibitory concentration (IC_{50}) $< 100 \mu\text{M}$, $n = 4$; measurements made in stratum radiatum). Conversely, compounds which have weak interactions with the NMDA receptor studied electrophysiologically^{1,4} (KA, QA, 2-amino-6-phosphohexanoate, L-APB) are poor displacers of NMDA-sensitive binding ($\text{IC}_{50} \geq 100 \mu\text{M}$, $n = 4$). This binding site may be the D-APV-sensitive, ³H-L-glutamate binding site described recently in rat brain membranes^{12,13}.

The NMDA-insensitive sites (Fig. 1b) are largely localized in the stratum lucidum of CA3, the commissural/associational layer of the dentate gyrus, and the pyramidal cell layer of CA1 and CA3. Binding in the former two layers was potentially displaced by low concentrations of KA (Fig. 1g). When KA was added in the absence of NMDA, the autoradiographic pattern in Fig. 1c was obtained. Further analysis revealed a K_i value for KA in the stratum lucidum of $67 \pm 5 \text{ nM}$ ($n = 2$). This represents the first demonstration of high-affinity displacement by KA of ³H-L-glutamate binding. As the distribution of these sites is similar to that found for ³H-KA binding^{14–16}, these data suggest that this glutamate binding site is identical to the KA binding site. Consistent with this interpretation is the observation that QA, L-glutamate and KA (but not D-glutamate, L-aspartate, D-aspartate, NMDA, D-APV, L-APB or ibotenate) are potent displacers of stratum lucidum ³H-L-glutamate binding ($\text{IC}_{50} < 1 \mu\text{M}$, $n = 2–4$)^{17–19}.

The binding in the pyramidal cell layer was not potentially inhibited by KA. When tested at $1 \mu\text{M}$, QA and α -amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA), but not KA, displaced most of this binding. At this concentration AMPA did not cause significant displacement of KA-sensitive binding. Figure 1d shows that $10 \mu\text{M}$ L-serine-O-sulphate (L-SOS) similarly exhibits higher affinity for the KA/NMDA-insensitive binding site than the KA site.

In whole brain^{10,20} and hippocampal²¹ membrane preparations, most (~80%) glutamate binding sites are dependent on the presence of Cl[−] ions. These sites are further expressed by the presence of Ca²⁺ ions and interact selectively with L-APB^{10,12,21}. Thus, Ca²⁺ and Cl[−] ions were added in an attempt to visualize these receptors which have been extensively characterized in membrane preparations containing Cl[−] in the assay buffer^{8–12,21,22}.

Addition of Ca²⁺ and Cl[−] ions introduces a pharmacologically and anatomically distinct ³H-L-glutamate binding site. As shown in Fig. 1e, binding now appears uniform across the CA1 and dentate gyrus layers. This is due to an increase in binding in the dentate gyrus molecular layer and in the stratum lacunosum-moleculare. Binding over the stratum lucidum is decreased by Ca²⁺ ions, consistent with the observation of Beaumont *et al.*²³ that Ca²⁺ is a potent inhibitor of KA binding. This fourth site becomes readily apparent when NMDA or D-APV is added (Fig. 1f). Binding is now predominant in the stratum lacunosum-moleculare and in the dentate gyrus molecular layer. Thus, the

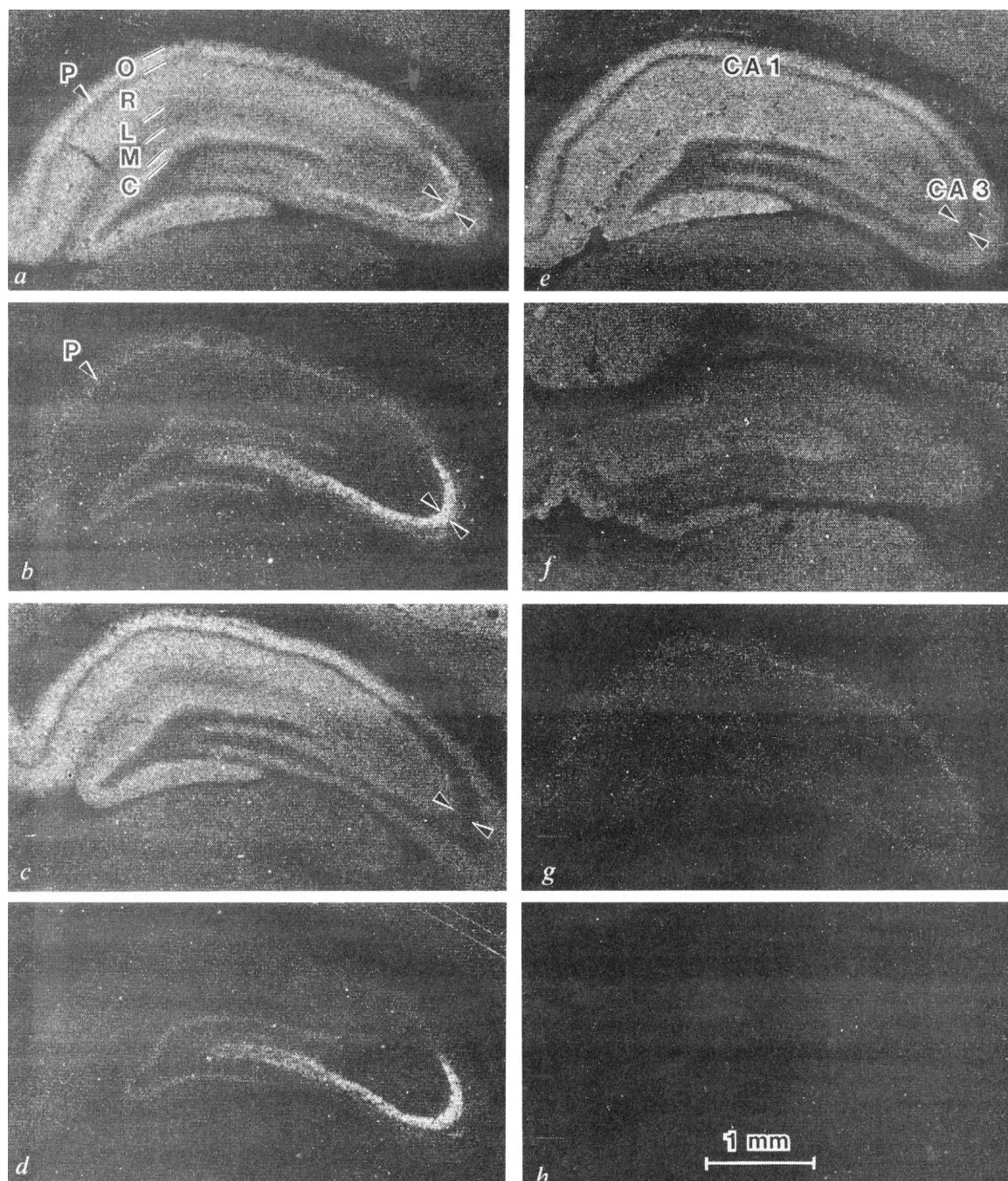


Fig. 1 Darkfield photomicrographs of ^3H -L-glutamate binding site autoradiograms (NTB-2 nuclear track emulsion) of rat hippocampus prepared in the presence of: *a*, buffer alone (50 mM Tris-acetate, pH 7.2); *b*, 100 μM NMDA; *c*, 10 μM KA; *d*, 100 μM NMDA plus 10 μM L-serino-*O*-sulphate; *e*, 2.5 mM Ca^{2+} and 20 mM Cl^- ; *f*, 100 μM NMDA, 2.5 mM Ca^{2+} and 20 mM Cl^- ; *g*, 100 μM NMDA and 1 μM KA; *h*, 0.5 mM L-glutamate. The paired arrows indicate the stratum lucidum; O, stratum oriens; R, stratum radiatum; L, stratum lacunosum-moleculare; P, pyramidal cell layer; M, outer two-thirds of molecular layer; and C, commissural/associational layer of the dentate gyrus. Fresh forebrains from Sprague-Dawley rats (45–60 days old) were removed, frozen and sectioned. The unfixed tissue sections were thaw-mounted on acid-washed microscope slides, then immediately (without drying) placed in ice-cold 50 mM Tris-acetate buffer, pH 7.2 for 2 h. After preincubation at 30 °C for 10 min, the sections were incubated for 30 min at 30 °C with 100 nM ^3H -L-glutamate (44.1 Ci mmol $^{-1}$; NEN), rinsed three times (30 s total) in ice-cold buffer and then air-dried⁴¹. Nonspecific binding (binding in the presence of 0.5 mM L-glutamate) represented 0–5% of the total binding. Radiolabelled tissue sections were placed in X-ray cassettes with LKB ^3H -Ultrafilm or clamped to NTB-2 (Eastman Kodak) emulsion-coated slides. Exposure time, 10–30 days. Quantitative autoradiography was performed on LKB film by including 13 ^3H -L-glutamate-brain paste standards⁴² per sheet of film. Density measurements were obtained using a De Anza image processing system with video camera⁴³. Each data point was the average of 50–2,000 areas within each layer for each of three or four sections. Inhibition constants¹⁰ and L-glutamate K_d s were calculated by Scatchard analysis of binding data obtained with increasing concentrations of unlabelled compound. K_d values were: 0.60 μM (radiatorium of CA1), 0.71 μM (stratum lucidum) and 0.74 μM (dentate gyrus molecular layer in the presence of 2.5 mM Ca^{2+} , 20 mM Cl^- and 100 μM NMDA—average values of duplicate determinations). In the absence of Ca^{2+} and Cl^- ions (see 1*a*) the total binding levels in pmol per mg protein were: R 0.49, L 0.40, M 0.43, C 0.45, stratum lucidum 0.46 ($n = 5$, s.d. < 15%). NMDA maximally displaced 80–90% of the binding in O, R, L and M layers, 75–80% of the binding in C and <10% in the stratum lacunosum-moleculare (0.20) and the dentate gyrus molecular layer (0.18). Due to the lower resolving capability of LKB film, we can only estimate pyramidal cell layer binding (see *g*) as being ~ 0.10 pmol per mg protein.

autoradiogram seen in Fig. 1b can be converted into that in Fig. 1f by adding Ca^{2+} and Cl^- ions. Of the compounds tested at 100 μM (NMDA, D-APV, L-APB, QA, KA, D- and L-glutamate, D- and L-aspartate, D- α -aminoadipate, L-homocysteate and ibotenate), only L-glutamate itself was a potent displacer of the $\text{Cl}^-/\text{Ca}^{2+}$ -enhanced glutamate binding.

As might be expected, the overall distribution of ^3H -L-glutamate binding (in the absence of Ca^{2+} and Cl^- ions) is similar to the distribution of ^3H -L-glutamate uptake²⁴, indicating an anatomical correspondence between uptake and receptor systems rather than suggesting that the binding observed here is due to uptake. Uptake site binding is thought to be Na^+ -dependent^{9,25,26}. In addition, the pharmacology of uptake inhibition is quite distinct, for example, NMDA, ibotenate and D- α -aminoadipate are poor inhibitors of uptake^{27,28}. Furthermore, binding is not thought to represent binding to enzyme sites as the K_m values for glutamate-related enzymes are in the millimolar range, and the overall distribution of the binding sites is quite pathway-specific.

Electrophysiological evidence suggests that both KA and NMDA receptors are involved in synaptic transmission^{29,30}. By using the *in vitro* spinal cord preparation, Watkins *et al.*^{29,30} demonstrated that the dorsal root-evoked-ventral root polysynaptic response is antagonized by compounds which selectively antagonize iontophoretically applied NMDA. In a preliminary account, the same group report that monosynaptic excitation of dorsal horn neurones is inhibited by a KA/QA-preferring antagonist, γ -D-glutamylaminomethylsulphonate. A synaptic function for KA binding sites is also suggested by our observations that KA binding sites are greatly enriched in purified synaptic junctions¹⁴.

Several lines of evidence suggest that glutamate may be the endogenous ligand for KA-sensitive glutamate binding sites within the stratum lucidum. Glutamate³¹, as well as glutamate immunoreactivity³² is found in high concentrations within this layer. Glutaminase-like immunoreactivity is found enriched within the mossy fibre boutons³³, and glutamate is taken up by mossy fibre boutons²⁴. Thus, endogenous glutamate interacting with the high density of KA-sensitive L-glutamate binding sites in the stratum lucidum may account for the observation that repeated activation of the mossy fibre pathway produces a KA-like lesion in the hippocampus³⁴. With the development of specific KA antagonists³⁰, the role of KA sites in mossy fibre synaptic transmission may soon be determined.

The termination fields of the Schaffer collaterals (CA1 stratum radiatum and stratum oriens) and the CA4-derived system (inner molecular layer of dentate gyrus) correspond to regions of high density for NMDA-sensitive ^3H -L-glutamate binding sites. Based on the uptake and release of acidic amino acids, both of these pathways are thought to use glutamate, aspartate or a related compound as their transmitter³⁵. Although NMDA receptors (which are antagonized by D-APV) are present in CA1 radiatum⁷, neither APV nor other NMDA antagonists block synaptic transmission at the Schaffer collateral terminal at concentrations $<1\text{ mM}$ ³⁶. Thus, the NMDA receptors in CA1 radiatum seem analogous to the high density of muscarinic cholinergic receptors also found in this layer which do not parallel cholinergic innervation³⁷. However, note that the excitotoxic action of quinolinic acid, which is somewhat selective for CA1, is antagonized by 2-amino-7-phosphonoheptanoic acid³⁸, an NMDA antagonist. Furthermore, based on electrophysiological experiments, it has been suggested that NMDA receptors within the stratum radiatum modulate synaptic transmission³⁹.

The QA/AMPA-sensitive, KA/NMDA-insensitive binding site possibly represents the QA receptor seen in electrophysiological studies. This is suggested by the observation that excitation induced by either QA or AMPA, but not KA or NMDA analogues, is antagonized by glutamic acid diethyl ester⁴⁰. The apparent selectivity of L-SOS has not been tested in electrophysiological experiments, but these results clearly predict a QA rather than KA-like action.

It is difficult to assess the significance of the $\text{Cl}^-/\text{Ca}^{2+}$ -enhanced L-glutamate binding site as only glutamate itself displayed a high affinity for the binding site. However, the distribution of these sites does correspond to the terminal fields of proposed glutamate-using pathways³⁵. Although the major class of glutamate binding sites observed in membrane preparations containing Cl^- ions are markedly dependent on the presence of Cl^- and Ca^{2+} ions^{10,20}, these sites are potentially inhibited by L-APB, ibotenate, L-aspartate and quisqualate¹² whereas the $\text{Cl}^-/\text{Ca}^{2+}$ -dependent binding observed in this study was not readily displaced by these compounds. APB-displaceable binding sites are freeze-labile¹³, which may account for their absence from this study.

Thus, we have identified four populations of ^3H -L-glutamate binding sites which differ in their anatomical distribution, pharmacological profile and ion sensitivity. Three of these populations have pharmacological characteristics similar to glutamate receptor classes described in electrophysiological studies. In other experiments we have found that the various sites have differing distributions in other brain regions as well as in the hippocampus (in preparation). In particular, the NMDA-sensitive site exhibits extensive anatomical specificity. Subdivisions are readily differentiated within cerebral cortex, basal ganglia, thalamus, hypothalamus, amygdala, septum, brain stem, spinal cord and other regions. In general, the distribution of this site is consistent with that of allocortical, corticofugal and primary sensory terminal fields, as would be expected for a glutamate system¹. The data demonstrate that simultaneously analysing the anatomical and pharmacological specificity of glutamate binding sites may resolve the complexity of glutamate receptor subtypes and yield valuable markers for pathways which use excitatory amino acids as neurotransmitters.

This work was supported by grants DRXRO-CB-19016-L (ARO) and RRO 1192 (NIH, image processing system).

Note added in proof: Autoradiographic visualization of the major ^3H -L-GLU binding site found in membrane preparations has recently been reported^{44,45}. The extent that the binding sites observed in this study are contributing to the binding observed by others is not known.

Received 23 May; accepted 18 August 1983.

- Watkins, J. C. & Evans, R. H. *A. Rev. Pharmac. Tox.* **21**, 165-204 (1981).
- Cotman, C. W., Foster, A. C. & Lanthorn, T. in *Glutamate as a Neurotransmitter* (eds DiChiara, G. & Gessa, G. L.) 1-27 (Raven, New York, 1981).
- Koerner, J. F. & Cotman, C. W. *Brain Res.* **216**, 192-198 (1981).
- Evans, R. H., Francis, A. A., Jones, A. W., Smith, D. A. S. & Watkins, J. C. *Br. J. Pharmac.* **75**, 65-75 (1982).
- Slaughter, M. M. & Miller, R. F. *Science* **211**, 182-185 (1981).
- Robinson, J. H. & Deadwyler, S. A. *Brain Res.* **221**, 117-127 (1981).
- Collinridge, G. L., Kehl, S. J. & McLennan, H. J. *Physiol. Lond.* **334**, 19-31 (1983).
- Foster, A. C. & Roberts, P. J. *J. Neurochem.* **31**, 1467-1477 (1978).
- Baudry, M. & Lynch, G. S. *J. Neurochem.* **36**, 811-820 (1981).
- Fagg, G. E., Foster, A. C., Mena, E. E. & Cotman, C. W. *J. Neurosci.* **2**, 958-965 (1982).
- Werling, L. L. & Nadler, J. V. *J. Neurochem.* **38**, 1050-1062 (1982).
- Fagg, G. E., Foster, A. C., Mena, E. E. & Cotman, C. W. *Eur. J. Pharmac.* **88**, 105-110 (1983).
- Fagg, G. E., Mena, E. E., Monaghan, D. T. & Cotman, C. W. *Neurosci. Lett.* **38**, 157-162 (1983).
- Foster, A. C., Mena, E. E., Monaghan, D. T. & Cotman, C. W. *Nature* **289**, 73-75 (1981).
- Monaghan, D. T. & Cotman, C. W. *Brain Res.* **252**, 91-100 (1982).
- Unnerstall, J. R. & Wamsley, J. K. *Eur. J. Pharmac.* **86**, 361-371 (1983).
- Simon, J. R., Contrera, J. F. & Kuhar, M. J. *J. Neurochem.* **26**, 141-147 (1976).
- London, E. D. & Coyle, J. T. *Molec. Pharmac.* **15**, 492-505 (1979).
- Slevin, J. T., Collins, J. F. & Coyle, J. T. *Brain Res.* **265**, 169-172 (1983).
- Mena, E. E., Fagg, G. E. & Cotman, C. W. *Brain Res.* **243**, 378-381 (1982).
- Whittemore, S. R., Mena, E. E., Monaghan, D. T. & Cotman, C. W. *Brain Res.* (in the press).
- Monaghan, D. T., McMillan, M. C., Chamberlain, A. R. & Cotman, C. W. *Brain Res.* (in the press).
- Beaumont, K. *et al. Life Sci.* **24**, 809-816 (1979).
- Storm-Mathisen, J. & Iversen, L. L. *Neuroscience* **4**, 1237-1253 (1979).
- Roberts, P. J. *Nature* **252**, 399-401 (1974).
- Vincent, S. R. & McGeer, E. G. *Brain Res.* **184**, 99-108 (1980).
- Balcar, V. J. & Johnston, G. A. R. *J. Neurochem.* **19**, 2657-2666 (1972).
- Roberts, P. J. & Watkins, J. C. *Brain Res.* **85**, 120-125 (1975).
- Evans, R. H., Francis, A. A., Hunt, K., Oakes, D. J. & Watkins, J. C. *Br. J. Pharmac.* **67**, 591-603 (1979).
- Davies, J. & Watkins, J. C. *J. Physiol., Lond.* **332**, 108P-109P (1982).
- Nitsch, C. in *Amino Acid Neurotransmitters* (eds DeFeudis, F. V. & Mandel, P.) 97-104 (Raven, New York, 1981).
- Storm-Mathisen, J. *et al. Nature* **301**, 517-520 (1983).
- Althuler, R. A. *et al. Neurosci. Abstr.* **8**, 663 (1982).
- Sloviter, R. S. & Damiano, B. P. *Neurosci. Lett.* **24**, 279-284 (1981).
- Cotman, C. W. & Nadler, J. V. in *Glutamate: Transmitter in the Central Nervous System* (eds Roberts, P. J., Storm-Mathisen, J. & Johnston, G. A. R.) 117-154 (Wiley, New York, 1981).

36. Koerner, J. F. & Cotman, C. W. *Brain Res.* **251**, 105–115 (1982).
 37. Storm-Mathisen, J. *Prog. Neurobiol.* **8**, 119–181 (1977).
 38. Schwartz, R., Whetsell, W. O. Jr & Foster, A. C. in *Excitoxins* (eds Fuxe, K., Roberts, P. J. & Scharz, R.) (Pergamon, Oxford, in the press).
 39. Collingridge, G. L., Kehl, S. J. & McLennan, H. J. *Physiol. Lond.* **334**, 33–46 (1983).
 40. Krosgaard-Larsen, P., Honore, T., Hansen, J. J., Curtis, D. R. & Lodge, E. *Nature* **284**, 64–66 (1980).
 41. Young, W. S. & Kuhar, M. J. *Brain Res.* **179**, 255–270 (1979).
 42. Unnerstall, J. R., Niehoff, D. L., Kuhar, M. J. & Palacios, J. M. *J. Neurosci. Meth.* **6**, 59–73 (1982).
 43. Altar, C. A., Walter, R. J. Jr, Neve, K. A. & Marshall, J. F. *Neurosci. Abstr.* **9**, 1113 (1983).
 44. Greenamyre, J. T., Penny, J. B. & Young, A. B. *Neurosci. Abstr.* **9**, 264 (1983).
 45. Halpain, S., Parsons, B. & Rainbow, T. C. *Eur. J. Pharm.* **86**, 313–314 (1982).

A shared antigenic determinant between natural killer cells and nervous tissue

S. Schuller-Petrovic*, W. Gebhart*, H. Lassmann†, H. Rumpold‡ & D. Kraft‡

* Department of Dermatology II, † Neurological Institute and ‡ Institute of General and Experimental Pathology, University of Vienna, A-1090 Vienna, Austria

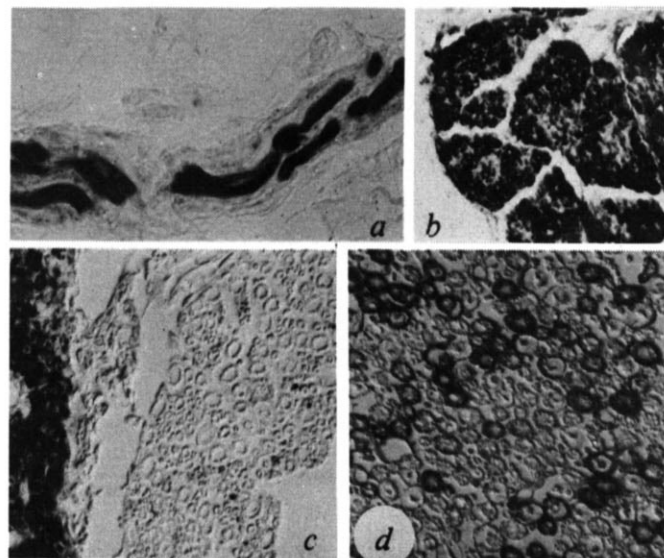


Fig. 1 Staining of peripheral nervous system components by anti-Leu-7 antibody. *a*, Small cutaneous nerve in human skin biopsy with intense staining of all myelin sheaths ($\times 250$). *b*, Anterior root of normal human spinal cord obtained from autopsy with intense staining of all myelin sheaths. Interference contrast ($\times 153$). *c*, Guinea pig spinal cord; peripheral nerve fibres in the anterior root are unstained. Intense staining of central myelin in the spinal cord (left margin of micrograph). (Interference contrast $\times 162$.) *d*, Anterior root of mouse spinal cord with intense staining of myelin in a selected portion of thick nerve fibres. (Interference contrast $\times 162$.) The tissue was fixed in 4% buffered formaldehyde and embedded in paraffin. Sections were deparaffinized with xylene/alcohol and stained with PAP-technique¹⁰. The following steps were used: (1) hydrogen peroxide/methanol; (2) normal swine serum (undiluted); (3) anti-Leu-7 antibody (Becton & Dickinson) 1:25 in Tris buffer with 10% normal swine serum; (4) rabbit anti-mouse immunoglobulins (Dako) 1:50; (5) swine anti-rabbit immunoglobulins (Dako) 1:40; (6) rabbit PAP-complex (Dako) 1:50; (7) diaminobenzidine reagent.

Considerable evidence for shared antigenic determinants between nervous elements and lymphocytes has accumulated^{1–3}. It has also been suggested that this cross-recognition may be involved in the pathogenesis of human neurological diseases such as myasthenia gravis^{1,4} and multiple sclerosis². We report here evidence that a marker for natural killer (NK) cells, anti-Leu-7 (HNK-1)⁵, specifically binds to components of human and rodent central nervous tissue as well as peripheral nervous tissue, especially to myelin sheaths. In contrast, another NK-cell marker (VEP13)⁶ did not react with nervous tissue. Since NK-cell function is impaired in a population of multiple sclerosis patients^{7–9}, the observed cross-reactivity indicates that autoimmunization against myelin may simultaneously cause a defect of NK-cell function. Furthermore, the shared antigenic determinant may help to identify a hitherto undefined nervous tissue antigen and simultaneously increase the knowledge about the nature of NK-cell antigens.

Immunocytochemistry was performed with an unlabelled immunoenzyme technique (PAP method)¹⁰ on paraffin sections. Tissue samples tested included biopsies of human skin (normal skin, lichen planus, halo nevus, neurofibroma), autopsy specimens of normal human spinal cord as well as normal guinea pig, rat and mouse spinal cord. All spinal cord samples contained portions of anterior and posterior spinal roots for determination of the staining patterns in the peripheral nervous system (PNS).

Anti-Leu-7 and VEP13 are both mouse monoclonal IgM antibodies recognizing human NK cells of peripheral blood and lymphatic tissue. Their specificity has been extensively characterized and it was shown that they react with different antigenic determinants on the NK-cell surface^{5,6,11–13}. In paraffin sections both antisera stained a small number of inflammatory cells, structurally comparable to large granular lymphocytes. No other structures of human and rodent central nervous system (CNS) or PNS and skin were recognized by VEP13 (Fig. 2g). Similarly, no non-specific staining of nervous elements was found with another monoclonal mouse IgM antibody directed against Ag₁ of *alternaria tenuis* or when normal mouse serum in dilutions

of 1:10 to 1:200 was used as primary layer. Nonspecific Fc-binding of immunoglobulins to myelin may lead to false-positive results in immunohistochemistry on cryostat sections of fresh frozen material. However, on paraffin sections of formaldehyde- or glutaraldehyde-fixed tissue, nonspecific Fc-binding of immunoglobulins to nervous tissue does not occur^{14–17}.

In all spinal cord samples from mice and rats, but not in those from humans and guinea pigs (irrespective of the primary layer used in the experiments) weak extracellular positive reaction was found in the peripheral portions of the spinal roots. This appears to be due to the recognition of mouse and rat immunoglobulin by the secondary antibody (rabbit anti-mouse immunoglobulin). Although immunoglobulins are generally excluded from the normal brain by the blood-brain barrier, leakage of serum proteins occurs in the peripheral portions of spinal roots and ganglia.

In contrast to the other monoclonal antibodies tested, anti-Leu-7 specifically stained several different components of the

Table 1 Nervous system antigens recognized by anti-Leu-7 monoclonal antibody

Species	Central nervous system							Peripheral nervous system				
	NC	AX	MY	OL	AG	EP	V,CT	NC	AX	MY	SC	V,CT
Human	(+)	–	+	+	–	+	–	ND	–	+	+	–
Guinea pig	(+)	–	+	+	–	–	–	(+)	–	–	–	–
Rat	(+)	–	+	+	–	–	–	ND	–	–	–	–
Mouse	(+)	–	+	+	–	–	–	ND	–	+	+	–

NC, nerve cells; AX, axon; MY, myelin; OL, oligodendroglia; AG, astroglia; EP, ependyma; V,CT, blood vessels, connective tissue elements, meninges; SC, Schwann cells. (+), Positive only after prolonged xylene/alcohol deparaffinization or delipidation with acetone or chloroform/methanol. ND, not determined.

* Myelin in mouse peripheral nerve recognized only in a small number of thick nerve fibres, mainly located in the anterior roots.

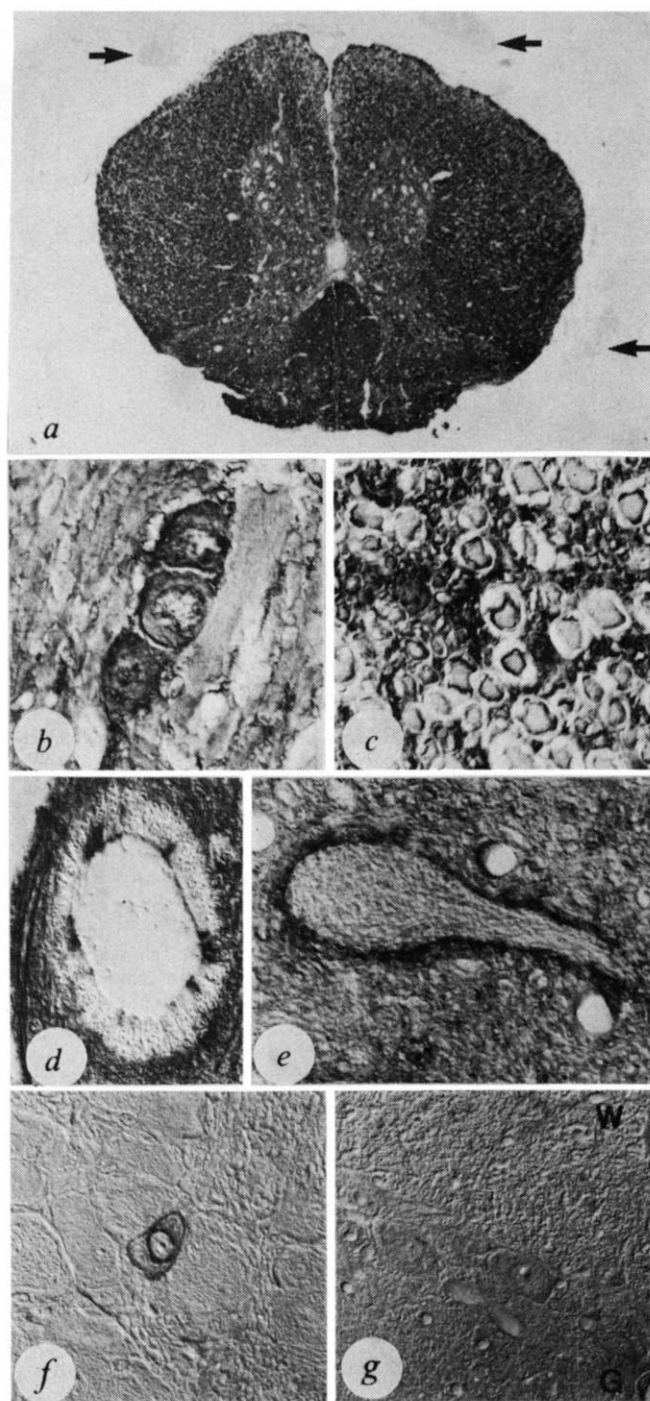


Fig. 2 Distribution of Leu-7-positive antigen in the nervous system. The immunohistochemical technique was identical to Fig. 1. *a*, Guinea pig spinal cord with intense reaction in the central nervous system, especially in the white matter. The spinal roots (arrows) are negative ($\times 61$). *b*, Longitudinal section of guinea pig spinal cord, fixed with 0.5% formaldehyde and 1.5% glutaraldehyde. The surface of interfascicular oligodendroglia shows binding of anti-Leu-7 antibody. Nuclei of the cells are stained with haematoxylin. Identification of oligodendrocytes was based on light-microscopic criteria of these cells, that is, the presence of rows of small round cells with dark round nuclei located in interfascicular position (interfascicular oligodendroglia) and by the presence of small delicate cell processes extending between myelin sheets. The oligodendroglia nature of the stained cells was further confirmed by staining of adjacent serial sections with Weil-Davenport silver impregnation technique. (Interference contrast $\times 612$.) *c*, Guinea pig spinal cord white matter. The small myelin sheaths are preferentially stained. In large fibres staining is mainly found in the inner and outer portions of the sheaths. (Interference contrast $\times 612$.) *d*, Central canal of human spinal cord; the apical cytoplasm and the luminal surface of some ependymal cells is stained by anti-Leu-7 antibody. (Interference contrast $\times 250$.) *e*, Anterior horn of guinea pig spinal cord; staining of perineuronal cell processes around perikaryon and major dendrite of a large spinal motoneurone. (Interference contrast $\times 612$.) *f*, Trigeminal ganglion of guinea pig. A small nerve cell shows Leu-7 antigen on the cell membrane and in perinuclear cytoplasm. (Interference contrast $\times 250$.) *g*, Section of guinea pig spinal cord reacted for VEP13 instead of Leu-7 antigen. VEP13 antibody was used in a dilution of 1:10. No reaction in the white (W) or grey matter (G) of the spinal cord. (Interference contrast $\times 250$.)

Table 2 Effect of tissue processing on recognition of Leu-7-positive nervous system antigens

Tissue processing	Effect
Paraffin embedding (formaldehyde fixation)	
Xylene/alcohol deparaffinization	Positive
Delipidation with acetone	Augmentation of reaction
Delipidation with chloroform/methanol	Augmentation of reaction
Detergent treatment (Tween 20)	Augmentation of reaction
Predigestion with pronase or trypsin	No effect on staining reaction
Paraffin embedding (glutaraldehyde fix)	
Xylene/alcohol deparaffinization	Positive (better structural preservation)
Delipidation with chloroform/methanol	Augmentation of reaction
Detergent treatment (Tween 20)	Augmentation of reaction
Paraffin embedding (human autopsy, tissue, formaldehyde fixation)	
Xylene/alcohol deparaffinization	Positive
Delipidation with chloroform/methanol	Augmentation of reaction

central and peripheral nervous system (Table 1). In humans all myelin sheaths in cutaneous nerves of the skin and in spinal roots were distinctly labelled by anti-Leu-7 antibody (Fig. 1*a*, *b*). In the CNS specific staining was found in central myelin (Fig. 2*a*, *c*), on the surface of oligodendrocytes (Fig. 2*b*), on the luminal surface of ependymal cells (Fig. 2*d*) and in perineuronal cell processes (Fig. 2*e*). The majority of nerve cells were negative, although in few neurones anti-Leu-7-positive labelling was noted on the cell surface and in the cytoplasm (Fig. 2*f*). In guinea pigs, rats and mice Leu-7 antigen was distributed in the CNS in a pattern similar to that in humans, except for negative staining of ependymal cells. In the peripheral nervous system, however, myelin and Schwann cells were negative in guinea pigs and rats (Fig. 1*c*). In mice a proportion of thick myelinated nerve fibres, predominantly located in the anterior roots, were specifically stained with anti-Leu-7 antibody (Fig. 1*d*).

Positive staining of nervous elements was seen with anti-Leu-7 antibody (Becton & Dickinson) at concentrations of up to

$2 \mu\text{g ml}^{-1}$. The Leu-7 antigen in the central and peripheral nervous system was relatively resistant to tissue processing (Table 2). Fixation with 4% buffered paraformaldehyde or with a mixture of 0.5% paraformaldehyde and 1.5% glutaraldehyde, embedding of tissue in paraffin, predigestion of the sections with pronase or trypsin or autolysis for 8 h did not destroy the antigenicity. However, delipidation augmented the immunohistochemical detection of the Leu-7 antigen and exposed additional antigenic determinants (Table 1).

The augmentation of the staining reaction by delipidation of the sections indicates that the antigen recognized by anti-Leu-7 is not a glycolipid antigen like galactocerebroside, gangliosides or sulphatide, since these antigens are removed from the sections by chloroform/methanol treatment. It is thus suggested that the antigenic determinant shared between NK cells and CNS tissue is present in cell membranes, partially buried in the lipid bilayer.

The chemical nature of the antigen recognized by the monoclonal anti-Leu7 antibody is at present not determined. The

immunohistochemical distribution in the CNS and PNS in the tested animal species is different from all other nervous system antigens investigated by similar techniques, which include myelin basic protein^{15,16}, myelin associated glycoprotein^{14,16} as well as P₀, P₁ and P₂ protein of peripheral myelin¹⁷. Thus anti-Leu-7 antibody apparently reacts with a yet undetermined membrane antigen, present mainly in central and peripheral myelin and to a smaller degree in nerve cell membranes.

The shared antigenic determinants between NK cells and myelin may be relevant in the pathogenesis of human inflammatory demyelinating diseases, especially of multiple sclerosis. An immune reaction against myelin, which is present in multiple sclerosis patients, may also destroy NK cells or functionally impair NK cell activity. This may be one possible explanation for the disturbance of NK-cell function, which has been found in a proportion of multiple sclerosis patients. Since a part of anti-Leu-7⁺ cells express the T3 as well as the T8

antigen^{11,12} and anti-Leu-7⁺ cells can also mediate suppressor cell function¹⁸, other lymphocyte subpopulations may also be affected, leading to further imbalance of immune regulation.

A similar concept has recently been put forward by Oger *et al.*² who were able to show a shared antigenic determinant between suppressor cells (T8⁺ cells) and ovine oligodendroglia. However, later studies revealed that the T8 antigen is not present on the surface of human, calf and rat oligodendrocytes^{19,20}. In this respect it appears especially important that in contrast to T8 antigen, Leu-7 antigen is shared by human CNS tissue and a subpopulation of human lymphocytes.

We thank Mrs M. Leiszer and Mrs U. Juszczak for technical assistance, and Mrs I. Lubenik for typing the manuscript. This study was partly supported by Science Research Funds, Austria, Project S25-07, and by Fonds Österreichische Krebsforschungsinstitute.

Received 19 May; accepted 2 September 1983.

1. Fuchs, S., Schmidt-Hopfield, I., Tridente, G. & Tarrab-Mazdai, R. *Nature* **287**, 162-164 (1980).
2. Oger, J., Szuchet, S., Antel, J. & Arnason, B. G. W. *Nature* **295**, 66-68 (1982).
3. Garson, J. A., Beverley, P. C. L., Coakham, H. R. & Harper, G. I. *Nature* **298**, 375-377 (1982).
4. Weiner, H. L. & Hauser, S. L. *Ann. Neurol.* **11**, 437-449 (1982).
5. Abo, T. & Balch, C. M. *J. Immun.* **127**, 1024-1029 (1981).
6. Rumpold, H., Kraft, D., Obexer, G., Böck, G. & Gebhart, W. *J. Immun.* **129**, 1458-1464 (1982).
7. Benczur, M. *et al. Clin. exp. Immun.* **39**, 657-662 (1980).
8. Merrill, J., Jondal, M., Seeley, J., Ullberg, M. & Siden, A. *Clin. exp. Immun.* **47**, 419-430 (1982).
9. Hauser, S. L., Ault, K. A., Levin, M. J., Garovoy, M. R. & Weiner, H. L., *J. Immun.* **127**, 1114-1117 (1981).

10. Sternberger, L. A. *Immunocytochemistry* 2nd edn (Wiley, New York, 1979).
11. Abo, T., Cooper, M. D. & Balch, C. M. *J. Immun.* **129**, 1752-1757 (1982).
12. Abo, T. & Balch, C. M. *J. Immun.* **129**, 1758-1761 (1982).
13. Rumpold, H., Obexer, G., Möschl, P. & Kraft, D. in *Protides of the Biological Fluids* (ed. Peeters, H.) (Pergamon, Oxford, in the press).
14. Sternberger, N. H., Quarles, R. H., Itoyama, Y. & Webster, H. de F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1510-1514 (1979).
15. Sternberger, N. H., Itoyama, Y., Kies, M. W. & Webster, H. de F. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2521-2524 (1978).
16. Itoyama, Y. *et al. Ann. Neurol.* **7**, 167-177 (1980).
17. Trapp, B. D., McIntyre, L. J., Quarles, R. D., Sternberger, N. H. & Webster, H. de F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3552-3556 (1979).
18. Tilden, A. B., Abo, T. & Balch, C. M. *J. Immun.* **130**, 1171-1175 (1983).
19. Traugott, U., Reinherz, E. L. & Raine, C. S. *Science* **219**, 308-310 (1983).
20. Hirayama, M., Lisak, R. P., Kim, S. U., Pleasure, D. E. & Silberberg, D. H. *Nature* **301**, 152-154 (1983).

Radiofrequency destruction of the tuberoinfundibular region of hypothalamus permanently abrogates NK cell activity in mice

Guido Forni*, Mauro Bindoni†, Angela Santoni‡, Natale Belluardo†, Anna E. Marchese† & Mirella Giovarelli*

* Department of Microbiology, University of Turin, Via Santena 9, 10126-Turin, Italy

† Department of Physiology, University of Catania, Viale A. Doria 6, 95125-Catania, Italy

‡ Chair of Virology, University of Perugia, Via del Giochetto, 06100-Perugia, Italy

Natural killer (NK) cells have an important role in non-adaptive resistance to tumours and their metastatic spread *in vivo*¹⁻⁵. Maturation of NK cells and the intensity of their activity are affected by many endogenous and external factors, as well as by regulatory cells¹. The possibility that some effects of the central nervous system on tumour resistance are mediated via NK activity has also been suggested⁶. Destruction of the tuberoinfundibular region of the hypothalamus in rodents led to a significant increase in tumour growth⁷⁻⁹. We show here that destruction of its ventromedial, dorsomedial and arcuate nuclei persistently abrogates NK activity in mice. By contrast, cortical lesion and operative stress depress it partially, and for a brief period only. Abrogation is the result of a block of NK lineage maturation, causing a severe decrease in the number of large granular lymphocytes (LGL), a lymphocyte population associated with NK activity^{10,11}. Macrophage, B- and T-lymphocyte functions, however, are not significantly affected. Agents inducing NK-cell maturation or activation such as polyinosinic-polycytidylic acid (poly(I:C)), interferon (IFN) and interleukin-2 (IL-2) restore NK activity, and normalize the number of LGL.

Seven-week-old C57BL/6 (B6) male mice (H-2^b) kept under natural lighting at 24-25 °C and on a standard diet (Charles River) were anaesthetized with a mixture of ketamine hydrochloride (Bristol-Meyers) and absolute alcohol. The nerve tissue was destroyed by electrothermocoagulation with a radiofrequency oscillator (500 Hz, 0-25 V, 0-200 mA), using stainless-steel electrodes insulated with epoxylite except for a 500-µm length at the tip⁸ oriented in the brain with a Kopf stereotaxic apparatus by means of Lehman's coordinates¹². The indifferent electrode was a banana plug inserted in the rectum. The brain was fixed, embedded in paraffin, cut into 12-µm coronal sections and stained with haematoxylin and eosin to determine the extent of lesion (Fig. 1). Age-matched, untreated, sham-operated mice, and mice with cerebral hemisphere lesion were used as controls. NK activity¹³ was determined for 8 weeks, starting 7 days after treatment. Virtually total, persistent abrogation was found in mice with hypothalamic lesion. In sham-operated mice, and those with cerebral hemisphere lesion, a marked decrease was observed 7 days after surgery, although this was followed by full recovery after 14 and 21 days, respectively (Fig. 2). There were no significant differences between each group with regard to body weight, rectal temperature or the weight of the main endocrine organs. Spleen weight, cell yield, proportions of B and T lymphocytes (determined by staining with fluorescent anti-Thy 1.2 and anti-mouse immunoglobulin antibody, respectively), and percentage of latex beads ingesting cells were similar in each group (data not shown).

To elucidate some features of this immune deficit, we evaluated macrophage and lymphocyte functions (Table 1). Peritoneal resident cells were tested for natural, poly(I:C) or lipopolysaccharide (LPS) induced cytotoxicity¹⁴. Spleen cell pools were assayed for proliferative response and IFN release¹⁵ towards B- and T-lymphocyte mitogens. Spleen T lymphocytes passed through nylon wool columns¹⁶ were also tested in mixed leukocyte culture against DBA/2 (H-2^d) stimulator spleen cells¹⁵. The cytolytic activity of T lymphocytes thus generated was tested against ⁵¹Cr-labelled P815 mastocytoma cells (H-2^d). The results, as well as those of other tests done 1 and 6 weeks after surgery (data not shown), indicate that no significant change took place after hypothalamic lesion.

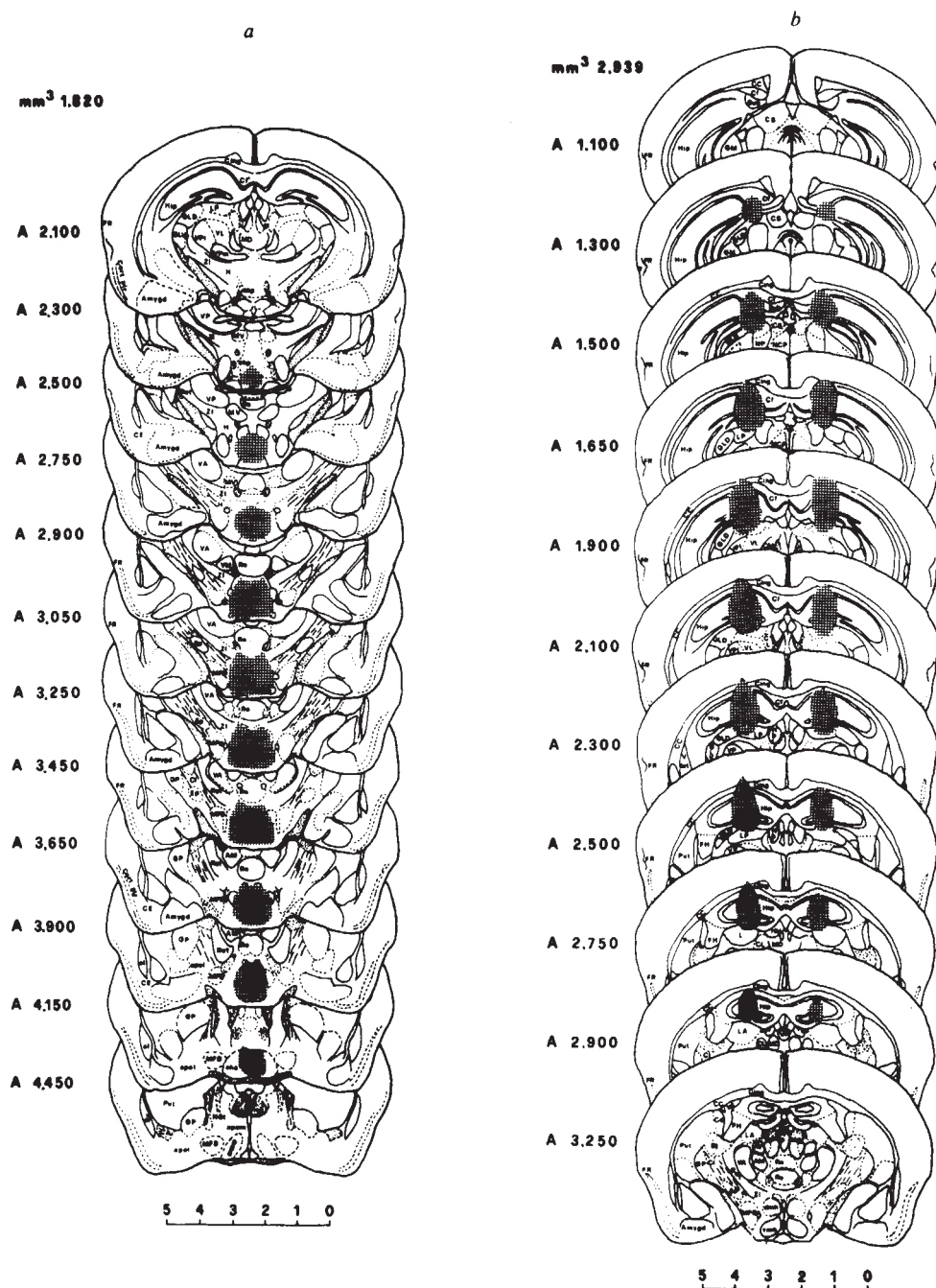


Fig. 1 Coronal brain sections showing the site and the size (hatched areas) of the hypothalamic lesions (bilateral destruction of ventromedial, dorsomedial, and arcuate nuclei between planes 2.500 and 3.500) (a), and hemisphere lesions (partial destruction of the hippocampus, dorsolateral nuclei of thalamus and corpus callosum between planes 1.300 and 3.050) (b). Volumes were calculated by conversion into three-dimensional terms as previously described⁸. This determination was performed on each mouse used in the experiment. The results obtained from mice bearing lesions outside these limits were discarded.

Abrogation of NK activity was not due to cell-mediated suppressor mechanisms. This was clear from mixture experiments between spleen cells from normal mice and mice with hypothalamic lesion, performed as previously described¹³ (data not shown). To determine whether abrogation of NK activity corresponded to a decrease in LGL, pools of spleen cells from untreated mice and mice with hypothalamic lesion were passed through nylon wool columns and fractionated by Percoll (Pharmacia) discontinuous density gradient centrifugation as described by Timonen *et al.*¹⁰ with slight modifications¹⁷. Spleens from lesioned mice yielded far fewer LGL than control spleens. The cytotoxicity pattern displayed by the cells in the Percoll fractions was strictly related to LGL distribution (Fig. 3).

As IFN markedly enhances NK activity and regulates LGL maturation¹⁷⁻¹⁹, we investigated its possible restoration by intraperitoneal (i.p.) injection of 5 mg per kg of poly(I:C), a potent inducer of IFN *in vivo*²⁰. Full restoration and LGL normalization were observed 36 h later (Fig. 3). The slightly different distribution of LGL and NK cytotoxicity in the Percoll fractions after poly(I:C) injection agrees with our recent findings¹⁷. NK activity in mice with hypothalamic lesion was also

enhanced by incubation of their spleen cells for 5 h with 10^3 IU of type I IFN, as described by Riccardi *et al.*¹⁹, or 24 h with 10 U ml⁻¹ of mitogen-free IL-2 as described by Kuribayashi *et al.*²¹ (data not shown).

We therefore conclude that abrogation of NK activity by hypothalamic lesion is attributable to drastic interference with NK cell maturation and activation and not to induction of suppressor cells. That it does not depend on substances released from the electrothermocoagulated nerve tissue is clear from the fact that destroyed tissue volume was always greater in the hemispheres than in the hypothalamus (Fig. 1).

Saxena *et al.*²² have shown that surgical ablation of the hypophysis results in a high degree of NK deficiency that could not be restored with IFN or poly(I:C). Moreover, ablation of the hypophysis in rodents greatly reduces spleen weight, lymphocyte yield and B- and T-lymphocyte reactivity²²⁻²⁴. In our study, the decrease in LGL number and NK activity cannot be a direct result of impaired hypophyseal function, because there is no sign of damage to this organ and its target glands. Hypothalamic lesion may, however, alter some of the functional connections between the infundibular region and the

Table 1 Functional reactivity of peritoneal resident and spleen cells

Effector cells from:	Spleen cell response against mitogens and alloantigens							
	% Specific lysis by peritoneal resident cells pretreated with			IFN release	³ H-Thymidine uptake			Cytotoxicity against
	None	Poly(I:C) (100 µg ml ⁻¹)	LPS (100 µg ml ⁻¹)		Con A (1.2 µg ml ⁻¹)	LPS (100 µg ml ⁻¹)	DBA/2 spleen	
Normal untreated mice	4.9*	41.0	38.7	52	79.0 (22)†	37.0 (17)†	54.9 (7)†	94
Sham-operated mice	6.0	36.8	39.0	51	76.5 (21)	35.1 (18)	ND	ND
Mice with cerebral hemisphere lesion	5.8	38.1	37.6	47	74.5 (22)	40.7 (18)	60.3 (7)	91
Mice with hypothalamic lesion	6.3	37.8	37.0	46	68.7 (21)	41.0 (16)	74.4 (8)	92

Effector cells were collected from 10 mice in each group. Peritoneal resident cells contained 70–75% monocyte-macrophage like cells, 4–7% polymorphonuclear cells, 20–25% lymphocytes as determined by differential cell counts made on stained smears prepared by cytocentrifugation. Suspensions (0.2 ml) containing 10^6 monocyte-macrophage-like cells per ml were placed in the wells of flat-bottomed microtitre plates. After 4 h incubation, the monolayers were shaken, vigorously washed to remove non-adherent cells¹⁴ and incubated overnight with poly(I:C), LPS or medium alone. Supernatants were aspirated and replaced with 0.2 ml of medium containing 2×10^4 ⁵¹Cr-labelled P815 mastocytoma cells per ml (50:1 effector:target, E:T). After 18 h incubation, 0.1 ml of the supernatant was collected and the specific ⁵¹Cr release was calculated¹⁴. Spleen cells (5×10^5 per well) were stimulated by concanavalin A (Con A) and *Escherichia coli* LPS in flat-bottomed microtitre plates and incubated for 72 h. IFN titre of the supernatants was assayed by the vesicular stomatitis virus plaque reduction assay, and expressed as the reciprocal of the supernatant dilution causing 50% reduction of plaques as compared with controls¹⁵. The spleen cell cultures and similar cultures of nylon wool enriched spleen T lymphocytes stimulated for 96 h by 2,000 rad irradiated DBA/2 spleen cells (1:1, R:S ratio) were also tested for proliferative response. 16 h before collection the cultures were pulsed with ³H-thymidine. Parallel cultures of T lymphocytes stimulated by DBA/2 spleen cells were collected after 5 days and assayed for cytotoxicity in a classic 4 h ⁵¹Cr-release test. The results are expressed as lytic units per 10^6 cells, the lytic unit being the number of effector cells required to cause 20% cytotoxicity. Virtually no cytotoxicity ($\leq 1\%$) was found against unrelated EL-4 (H-2^b) target cells. The s.d. of the means were constantly less than 10% for all the assays and are not shown. ND, not done.

* Specific ⁵¹Cr release. † Uptake expressed as the geometric mean of c.p.m. $\times 10^{-3}$, with stimulation index given in parentheses¹⁵.

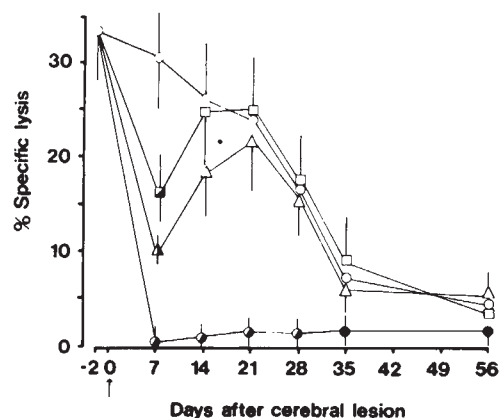


Fig. 2 Kinetics of NK activity against YAC-1 lymphoma target cells tested at 100:1 E:T ratio¹³ during the 8 weeks of the experiment. ○, Normal, untreated mice; □, sham-operated mice; △, mice with cerebral hemisphere lesion; ●, mice with hypothalamic lesion. Each point is the geometric mean for 6–10 mice independently tested + s.d. White and black symbols indicate values significantly different ($P < 0.01$) from corresponding values displayed by untreated controls, as determined by two-way analysis of variance. A group of 7-week-old mice with cytotoxicity levels between 30 and 35% specific lysis (not shown) was included in each determination. The experiment was performed independently three times with similar results.

hypophysis. Even so the consequences are selective for the NK lineage.

It is likely that hypothalamic lesion, directly, or through complex interactions with other encephalic areas, blocks the 'basal' production of neuroendocrine or immune system hormones, such as endorphin, IFN or IL-2, which may be essential for the complete maturation and activation of NK cells^{19,21,25}. This may be the case even though IFN production by peripheral T lymphocytes was normal following mitogen stimulation.

The present findings, although offering no explanation of the mechanisms involved in this regulatory block, provide a distinct cellular basis for a partial explanation of the altered resistance to tumours and infectious diseases that occurs during psychological stress, and after cerebral lesions in general and those of the hypothalamus in particular^{6-9,24,25}. Our experimental system is

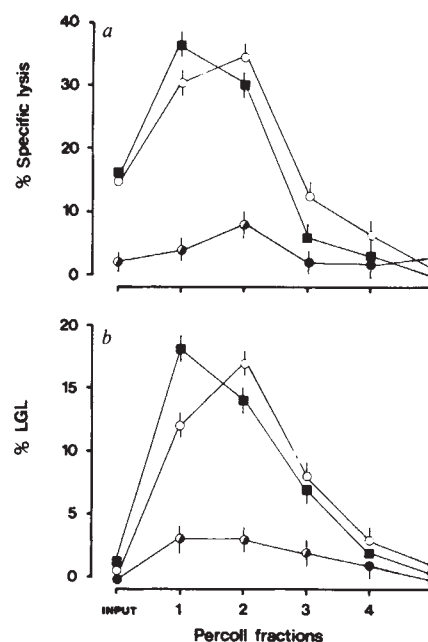


Fig. 3 Distribution of NK activity (a) and LGLs (b) after separation on Percoll gradients^{10,17}. Pools of spleen cells from 10 normal control mice (○), mice with hypothalamic lesion performed 3 weeks earlier (●) and lesioned mice injected intraperitoneally 36 h earlier with 5 mg per kg poly(I:C) (■) were passed through nylon wool columns and fractionated by discontinuous Percoll gradient centrifugation¹⁰, using seven Percoll concentrations ranging over 47.6–70.7%. 5×10^7 non-adherent spleen cells were placed on the top of the gradient and centrifuged at 300g for 40 min at room temperature. Each fraction was tested for NK activity against YAC-1 targets at 50:1 E:T, and examined for cell morphology¹⁰. Values represent the mean + s.d. of two independent experiments. White and black symbols indicate values significantly different ($P < 0.01$) from corresponding values displayed by untreated controls, as determined by two-way analysis of variance. Recovery from Percoll gradient was 80% of the original input. Viability was over 95%. Cell distribution was similar for all groups (fraction 1, $4 \pm 1\%$ of recovered cells; fraction 2, $15 \pm 2\%$; fraction 3, $28 \pm 3\%$; fraction 4, $12 \pm 4\%$; fractions 5 + F6, $4 \pm 2\%$).

also a simple model for investigation of the requirements for NK maturation *in vivo*.

We thank Dr R. B. Herberman and Dr C. Riccardi for advice and help. This work was supported by grants from the National Research Council (CNR), Italy, PFCCN 82.00236.96 and 82.00308.96.

Received 25 April; accepted 16 August 1983.

1. Herberman, R. B. (ed.) *NK Cells and Other Natural Effector Cells* (Academic, New York, 1982).
2. Riccardi, C., Puccetti, P., Santoni, A. & Herberman, R. B. *J. natn. Cancer. Inst.* **63**, 1041-1047 (1979).
3. Karre, D., Klein, G. O., Kiessling, R. & Roder, J. C. *Nature* **284**, 624-626 (1980).
4. Hanna, N. & Burton, R. C. *J. Immun.* **127**, 1754-1758 (1981).
5. Warner, J. F. & Dennert, G. *Nature* **300**, 31-34 (1980).
6. Herberman, R. B. in *Biological Mediators of Behaviour and Disease: Neoplasia* (ed. Levy, S.) (Elsevier, Amsterdam, in the press).
7. Bindoni, M., Belluardo, N., Licciardello, S., Marchese, A. E. & Cicirata, F. *Exptl Neurol.* **67**, 433-437 (1980).
8. Bindoni, M., Belluardo, N., Licciardello, S., Marchese, A. E. & Cicirata, F. *Neuroendocrinology* **30**, 88-93 (1980).
9. Bindoni, M., Belluardo, N., Marchese, A. E., Mudo', G. & Laguidara, A. *Neuroendocr. Lett.* **3**, 347-363 (1981).
10. Timonen, T., Saksela, E., Ranki, A. & Hayry, P. *Cell Immun.* **48**, 133-140 (1979).
11. Kumagai, K., Itoh, K., Suzuki, R., Hinuma, S. & Saitoh, F. *J. Immun.* **129**, 388-395d (1982).
12. Lehmann, A. *Atlas Stereotaxique du Cerveau de la Souris* (Centre National Recherche Scientifique, Paris, 1974).
13. Santoni, A., Riccardi, C., Sorci, V. & Herberman, R. B. *J. Immun.* **126**, 2329-2332 (1980).
14. Giovarelli, M., Landolfo, S., Whitmore, A. C. & Forni, G. *Eur. J. Cancer* **18**, 307-315 (1982).
15. Landolfo, S., Giovarelli, M., Viano, I. & Forni, G. *Immunogenetics* **9**, 245-253 (1979).
16. Julius, M., Simpson, E. & Herzenberg, L. W. *Eur. J. Immun.* **3**, 465-469 (1973).
17. Santoni, A., Herberman, R. B., Piccoli, M. & Ortaldo, J. *Fedn Proc.* **42**, 1380 (1983).
18. Trinchieri, G. & Santoli, D. *J. exp. Med.* **143**, 569-682 (1978).
19. Riccardi, C., Vose, B. M. & Herberman, R. B. in *NK Cells and Other Natural Effector cells* (ed. Herberman, R. B.) 909-915 (Academic, New York, 1982).
20. Djeu, J. Y., Heinbaugh, J. A., Holden, H. T. & Herberman, R. B. *J. Immun.* **122**, 175-181 (1979).
21. Kuribayashi, K., Gillis, S., Kern, D. E. & Henney, C. S. *J. Immun.* **126**, 2321-2326 (1981).
22. Saxena, O. B., Saxena, R. K. & Adler, W. H. *Int. Archs. Allergy appl. Immun.* **67**, 169-174 (1982).
23. Talwar, G. P. *et al. Recent Prog. Horm. Res.* **31**, 141-174 (1975).
24. Fabris, N., Pierpaoli, W. & Sorkin, E. *Clin. exp. Immun.* **9**, 227-240 (1971).
25. Mathews, P. M., Froelich, C. J., Sibbit, W. L. & Bankhurst, A. D. *J. Immun.* **130**, 1658-1662 (1983).
26. Solomon, G. F. & Amkkrut, A. A. *Rev. Microbiol.* **35**, 155-184 (1981).

Antibody-dependent cell-mediated antibacterial activity of intestinal lymphocytes with secretory IgA

A. Tagliabue*, L. Nencioni*, L. Villa*, D. F. Keren†, G. H. Lowell‡ & D. Boraschi*

* Laboratory of Immunopharmacology, Sclavo Research Center, 53100, Siena, Italy

† Department of Pathology, University of Michigan, Ann Arbor, Michigan 48109, USA

‡ Department of Bacterial Diseases, Gastroenterology and Immunology, Walter Reed Army Institute of Research, Washington, DC 20012, USA

Secretory antibodies of the IgA class (sIgA) are thought to have an important role in the defence against bacteria at mucosal surfaces—the level at which the infectious agents first come into contact with the host. However, the mechanism by which sIgA exert their antibacterial activity is still a matter of debate. After the recent discovery of receptors for the Fc portion of IgA (RFcα) on lymphocytes, monocytes and granulocytes of human, rabbit, guinea pig and mouse origin¹⁻⁶, it has been hypothesized that IgA also mediate antibody-dependent cellular cytotoxicity (ADCC). Indeed, ADCC mediated by human leukocytes against bacteria has been demonstrated in the presence of human circulating IgA⁷⁻⁹. As RFcα have also been shown to bind sIgA², we decided to investigate whether sIgA could mediate antibacterial ADCC when bound to lymphocytes from the murine gut-associated lymphoid tissues (GALT) which first interact with the invading bacteria. By using *Shigella* X16 (a hybrid strain between the enteric pathogen *Shigella flexneri* and *Escherichia coli*)¹⁰ as target in an *in vitro* assay that measures cell-mediated antibacterial responses¹¹, we

found that murine lymphocytes from GALT but not from other tissues are able to exert natural antibacterial activity against *Shigella* X16, and that sIgA significantly and specifically increase the natural antibacterial activity of GALT lymphocytes from mice and induce antibacterial activity in cells from the spleen, but not from the thymus or popliteal lymph nodes. Thus, we now propose a new role for sIgA in protecting the host against infectious agents at the mucosal level.

The intestinal secretions collected from chronically isolated rabbit ileal (Thiry-Vella) loops injected with *Shigella* X16 were used to obtain purified sIgA (<1.3 ng ml⁻¹ of IgG in preparations with 0.6 mg ml⁻¹ of sIgA)¹². These antibodies were then incubated for 2 h at 37°C with murine cells from different anatomical sites and *Shigella* X16. At the end of the incubation period, the pellets were vigorously resuspended (no bacteria adherent to leukocytes were evident by microscopic examination after this treatment) and appropriately diluted aliquots were plated on agar to assess colony-forming units (CFU)¹¹. Table 1 shows two representative experiments performed with different effector-to-target ratios and an optimal sub-agglutinating antibody dilution (1/500), as determined by the antibody titration reported in Table 2. The data in Tables 1 and 2 show that lymphocytes from the epithelium and the lamina propria of the small intestine, from Peyer's patches and from mesenteric lymph nodes, all exert natural antibacterial activity. This confirms and extends recent observations obtained using *Salmonella typhimurium* in this *in vitro* assay¹¹. Interestingly, the spontaneous antibacterial activity against *Shigella* X16 was peculiar to GALT lymphocytes, as cells from thymus, spleen or popliteal lymph nodes did not affect the bacterial growth by themselves. When purified anti-*Shigella* X16 sIgA were added to the *in vitro* assay, a significant increase in the antibacterial activity of GALT and spleen leukocytes was observed. The lack of effect of sIgA on cells from thymus and popliteal lymph nodes (Tables 1, 2), the absence of any increase in the anti-*Shigella* X16 activity with crude fluid from ileal loops of normal rabbits (data not shown), and the results of cross-reactivity experiments with *Salmonella tel aviv* (Table 3) reasonably rule out nonspecific effects in the system used. Moreover, the extremely low contamination with IgG of our preparation of sIgA¹² (the ratio IgG:sIgA was less than 2.1 × 10⁻⁶) makes it highly unlikely that IgG-dependent activity has a role in our system, particularly with antibody dilutions as high as 1/500. Finally, direct evidence that sIgA and not contaminating antibodies of other classes determined the activity against *Shigella* X16 was obtained using sIgA purified on affinity columns. In fact, sIgA recovered from the column still showed antibacterial activity when bound to lymphocytes from normal mice (expt 3, Table 1). Further, it must be stressed that lymphocytes from Peyer's patches and mesenteric lymph nodes cannot exert IgG-mediated ADCC against bacterial and non-bacterial targets^{11,13}. Thus, the sIgA-dependent antibacterial ADCC observed *in vitro* suggests a specialist function for the cells from these organs.

In view of the fact that only an incomplete reduction of bacterial viability was obtained, the ratios between effector cells and bacteria used in this *in vitro* assay (from 25:1 to 200:1) might seem too high, raising criticism as to their relevance to an *in vivo* situation. However, it must be considered that the *in vitro* doubling time of *Shigella* X16 is 20 min and that the effector:target ratio thus rapidly decreases during the 2 h assay. Furthermore, it is possible that only a minor fraction of the total population used would have antibacterial activity, as shown in many other assays measuring cell-mediated activities¹⁴, and this would bring the actual effector:target ratios to much lower values. Finally, it has been reported that 90-95% of patients having Gram-negative bacterial sepsis have bacteraemias of <1 × 10² bacteria per ml of peripheral blood¹⁵; thus, the *in vitro* system seems to resemble the *in vivo* pathological condition.

Previous studies^{16,17} have shown that the population of cells obtained from the intestinal mucosa after Percoll fractionation comprises essentially lymphocytes. Thus, at least at this level, the effector cell of both natural and sIgA-dependent antibac-

Table 1 Natural and sIgA-mediated antibacterial activity of intestinal and peripheral leukocytes against *Shigella* X16

Cell source	sIgA	Antibacterial index				Statistical analysis
		25*	50	100	200	
Expt 1 Spleen	—	—	-2	-7	-5	—
	1/500†	—	15	22	34	‡
Gut epithelium	—	—	13	22	36	—
	1/500	—	25	34	41	(3.5)§
Gut lamina propria	—	—	22	29	34	—
	1/500	—	40	44	45	(7.4)§
Peyer's patches	—	22	25	28	—	—
	1/500	31	41	45	—	(6.5)§
Expt 2 Spleen	—	—	0	-6	-2	—
	1/500	—	23	27	41	‡
Mesenteric lymph nodes	—	—	19	20	26	—
	1/500	—	33	38	42	(12.5)§
Popliteal lymph nodes	—	—	-4	-11	-9	—
	1/500	—	-6	-6	-4	NS
Peyer's patches	—	8	6	14	—	—
	1/500	15	20	25	—	(8.4)§
Thymus	—	—	1	2	6	—
	1/500	—	3	6	6	NS
Expt 3 Spleen	—	—	-3	-3	-5	—
	1/500	—	6	16	22	‡
	1/100	—	22	34	36	‡
Peyer's patches	—	13	20	27	—	—
	1/500	15	26	31	—	NS
	1/100	28	35	38	—	(3.4)§

Single cell suspensions were obtained from spleen, Peyer's patches and lymph nodes of 8-week-old C3H/HeN male mice by gentle teasing. Leukocytes from gut epithelium and lamina propria were obtained as described previously^{16,17} and purified on Percoll gradients¹⁷. Antibacterial activity^{7-9,11} was assessed as follows: 10^4 bacteria were placed in 15-ml conical tubes together with either diluted anti-*Shigella* X16 sIgA antibodies or medium (RPMI 1640 plus 10% fetal bovine serum) and centrifuged at 1,300g for 10 min at 4°C. Lymphocytes were then added and the tubes were again centrifuged at 500g for 5 min at 4°C. To maintain the optimal proportion of reactants, the final volume of the mixture was 0.3 ml. The experimental and control tubes (which contained bacteria but no cells) were then incubated at 37°C for 2 h, after which the pellets were vigorously resuspended, diluted and plated on Petri dishes with agar-tryptose. CFU were counted after overnight incubation. Duplicate tubes were set up for each experimental group and two Petri dishes were prepared for each tube. The percentage of antibacterial activity was expressed as an antibacterial index = $100 - [100 \times (\text{CFU of experimental tubes}) / (\text{CFU of control tubes without leukocytes})]$. sIgA alone had no effect on bacterial survival. Antibacterial activity was analysed statistically by parallel line assay²¹ after logarithmic transformation of the variables. Potency of experimental groups with sIgA versus no sIgA was estimated as the ratio between the number of cells of test and control groups giving equal antibacterial indices. Expts 1 and 2 are representative of 11 performed. The variability of the experimental system was low: for example, in 11 experiments the mean antibacterial indices $\pm$ s.e. for natural and antibody-mediated activity were, respectively, -7.5 ± 2.9 and 32.6 ± 2.1 for splenocytes at the 200:1 ratio, and 27.8 ± 3.0 and 43.3 ± 4.1 for leukocytes from Peyer's patches at the 100:1 ratio. Expt 3 was performed using sIgA purified in affinity chromatography columns containing anti-IgA antibodies (Miles, Elkhart) conjugated to CNBr-activated Sepharose 4B (Pharmacia).

* Effector:target bacteria (E:T) ratio.

† Antibody dilution.

‡ Control and test groups have significantly different slopes ($P \leq 0.05$).

§ Control and test groups have parallel slopes. Potency (in parentheses) is significant ($P \leq 0.05$). NS, not significant.

|| Purified by affinity chromatography.

Table 2 Titration of sIgA antibodies against *Shigella* X16

Cell source	Antibacterial index for sIgA diluted:					
	No IgA	1:250	1:500	1:1,000	1:2,000	1:4,000
Spleen*	0	4	-1	3	3	5
Gut epithelium*	5	14‡	37‡	30‡	14‡	4
Peyer's patches†	34	42	53‡	46‡	36	33
Mesenteric lymph nodes*	27	32	43‡	40‡	36‡	27
Popliteal lymph nodes*	30	36‡	46‡	42‡	36‡	30
Thymus*	2	2	2	2	4	3
	0	2	2	5	3	2

Experimental conditions as in Table 1.

* E:T ratio, 200:1; † E:T ratio, 100:1.

‡ $P \leq 0.05$ versus no sIgA as assessed by analysis of variance.

Table 3 Specific effect of sIgA against *Shigella* X16

Cell source	sIgA anti- <i>Shigella</i> X16*	IgA anti- <i>S. tel aviv</i> *	Antibacterial index against:	
			<i>Shigella</i> X16	<i>S. tel aviv</i>
Spleen†	—	—	-2	30
	—	+	4	59§
	+	—	40§	26
Peyer's patches‡	—	—	32	-28
	—	+	27	20§
	+	—	46§	-16

Experimental conditions were as described in Table 1. IgA antibodies against *S. tel aviv* (from our own bacterial collection) were obtained from ascitic fluid of BALB/c mice injected with MOPC 384 plasmacytoma²².

* Dilution 1:500; † E:T ratio 200:1; ‡ E:T ratio 100:1.

§ Significant difference ($P \leq 0.05$) compared with corresponding control without antibodies, assessed as in Table 2.

Indeed, an increasing number of reports suggest that, besides macrophage-dependent activities, cell-mediated responses such as natural killer (NK) activity and ADCC, at first defined as effector mechanisms against neoplastic cells, might also have a role against viruses, fungi, bacteria and protozoa¹⁸. As many infectious agents contact the host at the mucosal level, it might be expected that these immunosurveillance mechanisms should be especially active at these anatomical sites. It is striking that large granular lymphocytes (LGL)—the main effector cells of NK activity and ADCC against tumour cells in many species^{19,20}—are present in high numbers in the intestinal epithelium and lamina propria, and are still able to exert NK activity against tumour cells at this level^{16,17}. As NK and ADCC seem to be different functions of the same lymphocyte, that is, the LGL^{18,19}, it might be suggested that the gut LGL also have a role in the antibacterial sIgA-dependent ADCC.

Thus, our finding that GALT lymphocytes, and in particular intra-epithelial lymphocytes, exert natural and sIgA-dependent activity against an enteric pathogen further supports the hypothesis that cell-mediated responses of the type discussed above might be relevant as a first line of defence against infectious diseases. Apart from the obvious theoretical implications, these results could stimulate new approaches to obtain more rational and effective vaccines for gastrointestinal infections.

Received 12 May; accepted 18 August 1983.

1. Strober, W., Hague, N. E., Lum, L. G. & Henkart, P. A. *J. Immun.* **121**, 2440-2445 (1978).
2. Lum, L. G. *et al. J. Immun.* **122**, 65-69 (1979).
3. Lum, L. G., Muchmore, A. V., O'Connor, N., Strober, W. & Blaese, R. M. *J. Immun.* **123**, 714-719 (1979).
4. Fanger, M. W., Shen, L., Pugh, J. & Brenner, G. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3640-3644 (1980).
5. Stafford, H. A. & Fanger, M. W. *J. Immun.* **125**, 2461-2466 (1980).

6. Arnaud-Battandier, F., Hague, N. E., Lum, L. G., Elson, C. O. & Strober, W. *Cell. Immun.* **55**, 106–113 (1980).
7. Lowell, G. H., Smith, L. F., Artenstein, M. S., Nash, G. S. & MacDermott, R. P. *J. exp. Med.* **150**, 127–137 (1979).
8. Lowell, G. H. *et al.* *J. Immun.* **125**, 2778–2784 (1980).
9. Lowell, G. H., Smith, L. F., Griffiths, J. M. & Brandt, B. L. *J. exp. Med.* **152**, 452–457 (1980).
10. Formal, S. B., La Brec, E. H., Kent, T. H. & Falkow, S. *J. Bact.* **89**, 1374–1382 (1965).
11. Nencioni, L., Villa, L., Boraschi, D., Berti, B. & Tagliabue, A. *J. Immun.* **130**, 903–907 (1983).
12. Keren, D. F. *Infect. Immunity* **24**, 441–448 (1979).
13. Arnaud-Battandier, F., Bundy, B. M., O'Neill, M., Bienenstock, J. & Nelson, D. L. *J. Immun.* **121**, 1059–1065 (1978).
14. Herberman, R. B. (ed.) *Natural Cell-mediated Immunity Against Tumours* 1321 pp. (Academic, New York, 1980).
15. Kreger, B. E., Craven, D. E., Carling, P. C. & McCabe, W. R. *Am. J. Med.* **68**, 332–343 (1980).
16. Tagliabue, A., Luini, W., Soldateschi, D. & Boraschi, D. *Eur. J. Immun.* **11**, 919–922 (1981).
17. Tagliabue, A., Befus, A. D., Clark, D. A. & Bienenstock, J. *J. exp. Med.* **155**, 1785–1796 (1982).
18. Herberman, R. B. *Clin. Immun. Rev.* **1**, 1–65 (1981).
19. Timonen, T., Ortaldo, J. R. & Herberman, R. B. *J. exp. Med.* **153**, 569–582 (1981).
20. Luini, W., Boraschi, D., Alberti, S., Aleotti, A. & Tagliabue, A. *Immunology* **43**, 663–668 (1981).
21. Finney, D. J. *Statistical Methods in Biological Assay* 2nd edn, 99–139 (Griffin, London, 1971).
22. Potter, M. *Physiol. Rev.* **52**, 631–719 (1972).

Role for mouse macrophage IgG Fc receptor as ligand-dependent ion channel

John Ding-E Young, Jay C. Unkeless,
Tsen M. Young*, Alexander Mauro
& Zanvil A. Cohn

Department of Cell Physiology and Immunology,
The Rockefeller University, New York, New York 10021, USA
*Massachusetts Institute of Technology–Harvard University
Division of Health Sciences and Technology, Cambridge,
Massachusetts 02138, USA

The interaction of ligands with the mouse macrophage Fc receptor which binds IgG2b and IgG1 immune complexes (FcR_{γ2b/γ1}) triggers phagocytosis and secretion of various mediators of inflammation^{1–4}. FcR_{γ2b/γ1} has been purified using a monoclonal anti-FcR antibody, 2.4G2 (refs 5, 6), and seems to be an integral membrane glycoprotein of molecular weight (*M_r*) 47,000–60,000 (ref. 6). Monoclonal antibody 2.4G2 is suitable as a tool for functional studies of FcR because it binds to a functional site of the receptor and induces cellular responses that are normally associated with the occupied receptor^{5,7,8}. We reported previously that binding of ligands to the macrophage FcR resulted in Na⁺/K⁺ ion fluxes through the plasma membrane, and that similar ion fluxes were observed in proteoliposomes containing reconstituted FcR^{9,10}. We have now incorporated FcR into planar lipid bilayers and report here that FcR_{γ2b/γ1} forms ligand-dependent cation-selective ion channels, with a conductance of 60 pS in 1 M KCl and an average open channel lifetime of 250 ms. The conductance decays to baseline levels within a few minutes. These results suggest a receptor-ionophore model for the signalling of phagocytosis and inflammatory responses.

FcR_{γ2b/γ1} purified as described previously^{5,6}, was reconstituted into lipid vesicles by detergent dialysis or detergent dilution¹¹. The integration of FcR_{γ2b/γ1} into liposomes was confirmed by isopycnic centrifugation (Fig. 1). To assay directly for FcR_{γ2b/γ1} in gradient fractions, a monoclonal sandwich radioimmunoassay with ¹²⁵I-2.4G2 Fab as labelled probe was used¹².

FcR–proteoliposomes were incorporated into planar bilayers by transforming lipid monolayers containing FcR into stable bilayers¹³. Electrical measurements were made as elsewhere¹⁴. Monolayers were prepared by a detergent-dilution protocol^{15–17} (Fig. 2). The larger size of the liposomes (>5,000 Å) formed by detergent dilution comparison with those formed by detergent dialysis^{11,18} may facilitate the spontaneous formation of the lipid monolayer¹⁶.

Bilayers formed from monolayers of phosphatidylcholine–cholesterol (either protein-free or containing FcR up to 5 µg per membrane) by the detergent-dilution method showed baseline conductance levels of <10 pS and a capacitance of 0.72 ± 0.06 µF cm^{–2}. These values were not changed by addition of ligands to protein-free bilayers, demonstrating that the ligands *per se* have no effect on membrane conductance. Addition of ligands to the *trans*, or FcR-free, side of the membrane resulted in no or a small increase in membrane conductance (Fig. 2). By contrast, when ligands were added to the *cis* side, a dramatic increase in membrane conductance was observed, the magnitude of which depended on the concentrations of both the receptor and the ligand used (Fig. 2). This sidedness in response suggested that FcR oriented in a vectorial fashion during monolayer → bilayer assembly or that denaturation of the receptor with the 'wrong' orientation occurred at the air–water interphase. An increase in conductance was observed with voltages of both polarities.

The conductance increase could be activated by a variety of ligands specific for FcR (Fig. 2). Insoluble immune complexes gave a larger response than soluble complexes or 2.4G2 IgG, which was more effective than 2.4G2 Fab. Controls in which monomeric anti-dinitrophenyl DNP IgG was added to the *cis* side showed no conductance changes. These results indicate that the state of aggregation of FcR by ligands bears directly on the conductance changes observed. These observations mirror the response of intact macrophages to immune complexes in the triggering of release of mediators of inflammation¹⁹.

The ligand-dependent increase in conductance always declined to baseline levels after a few minutes and was not reversed by subsequent addition of excess ligand (Fig. 2a). This inactivation or desensitization suggests that the perturbation of the receptor by ligand results in transient conformational change of the receptor followed quickly by the relaxation of the receptor to a new unresponsive state that will not allow ions to pass through the membrane.

Single FcR-channel fluctuations were obtained by reducing the amount of receptor in the bilayer. Figure 3 shows two such experiments, in which we used 2.4G2 IgG as ligand. The conductance changes always occurred in discrete fluctuations, which were indicative of opening and closing of single or groups of single channels. The responses observed in KCl and NaCl were similar. With a time resolution of 1 ms, we could not resolve the risetime of individual steps. This represents a flux over 10⁶ ions per s per molecule, which would be much too fast for any type of transport mechanism other than an ion-channel mechanism^{20,21}.

To resolve single-channel fluctuations, the ionic strength of the bathing buffer was increased (Fig. 4). The conductance per channel was 60 ± 5 pS (1 M KCl) and 54 ± 6 pS (1 M NaCl) and the average lifetime of the open channel (within the first 5 min of addition of ligands) was 250 ms. No intermediate states were observed. Single-channel fluctuations were difficult to maintain because of the spontaneous inactivation of the conductance over time, the frequency of the open state decaying in an exponential fashion after the first few minutes of contact with ligand. In 1 M KCl, the current–voltage (*I*–*V*) plot for single channels was linear, indicating ohmic behaviour (Fig. 5). Thus, the inactivation or desensitization of the FcR_{γ2b/γ1} is probably not due to a variable conductance of the channel itself but rather to a decreased frequency of opening of the bound receptor–channel. The kinetics of the channel depended on the lipid composition (and thus fluidity) of the bilayer, the open lifetime being shorter when lipids extracted from macrophages were used to form the bilayer.

The ion selectivity of the FcR–channel triggered by 2.4G2 IgG was determined from the reversal potential necessary to null current flow from a 10-fold higher KCl or NaCl concentration in the *cis* compartment and was calculated using the Nernst–Planck relationship¹⁷. Because of the rapid spontaneous inactivation of conductance increase following addition of ligands, such experiments could not be performed on a

Fig. 1 Flotation of FcR-proteoliposomes through sucrose gradients. Liposomes (a-e) were prepared from egg phosphatidylcholine and cholesterol (3 M:1 M), mixed with buffer, 25 mM β -D-octylglucoside and FcR (protein/lipid = 1:10 by wt), from which the detergent was dialysed away. Alternatively (f), β -D-octylglucoside (25 mM), lipids and Fc receptor (in the same ratio) were diluted 30 $\times$ with buffer. The buffer (A) used was 0.15 M NaCl, 4 mM $MgCl_2$, 3 mM $CaCl_2$, 3 mM NaN_3 , 10 mM HEPES, pH 7.4. Aliquots of FcR-proteoliposomes (5 μ g in 300 μ l) in 40% sucrose (wt/vol) were transferred to cellulose nitrate tubes, overlaid with 5 ml of a continuous sucrose gradient (5-35%) and centrifuged for 16 h at 190,000g at 4 $^{\circ}C$. Fractions (0.25 ml) were collected and numbered from the top of the gradient. Total radioactivity per fraction was given (except in d). Purified FcR was labelled with ^{125}I

as described previously⁶ and used in trace amounts as a marker (input 2×10^4 c.p.m.). [Choline-Me- ^{14}C]phosphatidylcholine (4×10^4 c.p.m.) was used as a marker for lipid. Sucrose density was determined by refractometry. In a, vesicles were formed in the absence of protein; vesicles floated to the top of the gradient. b, Lipid was absent during reconstitution, and FcR was found in the bottom of the tube. c, Vesicles containing FcR and lipid. d, The amount of FcR in each fraction of an experiment similar to c (with 1 μ g protein/10 μ g lipid) was directly assayed in the absence of radiolabelled markers by a monoclonal sandwich radioimmunoassay¹² with ^{125}I -2.4G2 Fab (input 0.25 μ g ml $^{-1}$, at 3.7×10^3 c.p.m. ng $^{-1}$) as the measuring probe. The c.p.m. of ^{125}I -Fab-FcR complex for aliquots of 20 μ l are given. The profile for FcR coincided with that obtained using tracer ^{125}I -FcR and ^{14}C -phosphatidylcholine (c). e, Reconstitution as in c in the presence of 250 μ g bovine albumin (final vol. 300 μ l). The shift in isopycnic density indicates that proteoliposomes formed trapped albumin and thus are more dense. f, Proteoliposomes (5 μ g protein) formed by detergent dilution were pelleted first by centrifugation at 100,000g for 90 min, then subjected to a flotation gradient as before.

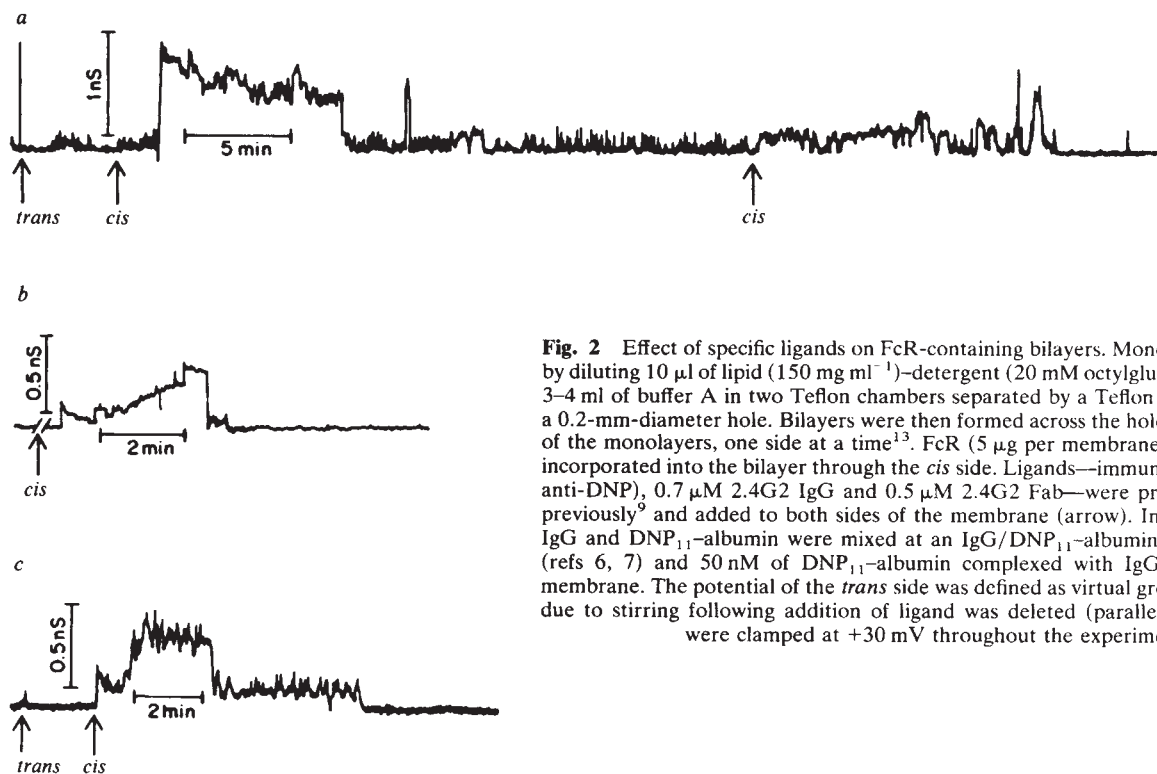
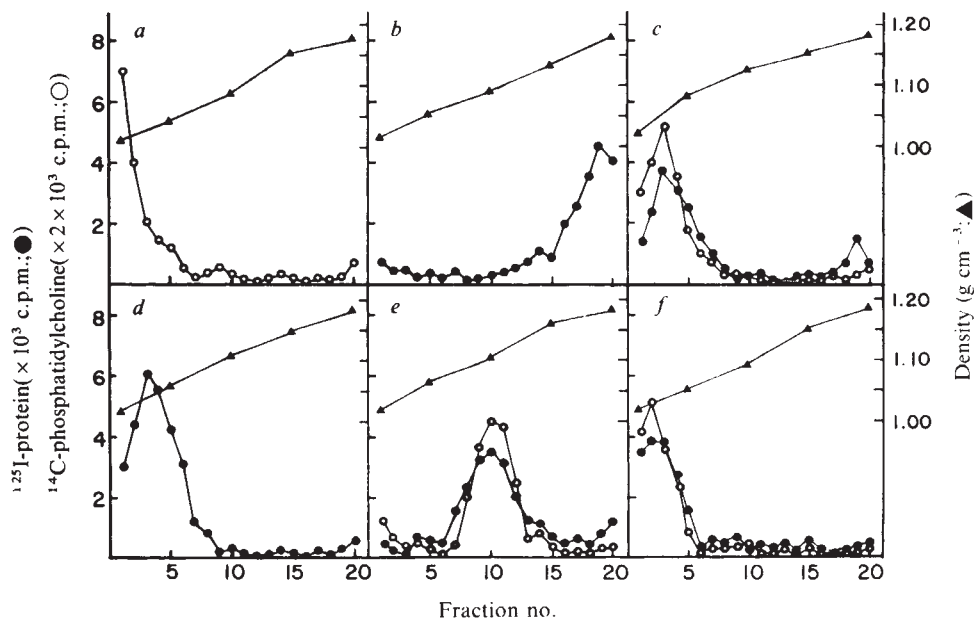


Fig. 2 Effect of specific ligands on FcR-containing bilayers. Monolayers were formed by diluting 10 μ l of lipid (150 mg ml $^{-1}$)-detergent (20 mM octylglucoside) mixture with 3-4 ml of buffer A in two Teflon chambers separated by a Teflon partition containing a 0.2-mm-diameter hole. Bilayers were then formed across the hole by raising the level of the monolayers, one side at a time¹³. FcR (5 μ g per membrane or 2×10^{-8} M) was incorporated into the bilayer through the cis side. Ligands—immune complexes (DNP-anti-DNP), 0.7 μ M 2.4G2 IgG and 0.5 μ M 2.4G2 Fab—were prepared as described previously⁹ and added to both sides of the membrane (arrow). In c, rabbit anti-DNP IgG and DNP₁₁-albumin were mixed at an IgG/DNP₁₁-albumin molar ratio of 6:1 (refs 6, 7) and 50 nM of DNP₁₁-albumin complexed with IgG was added to the membrane. The potential of the trans side was defined as virtual ground. In b, the noise due to stirring following addition of ligand was deleted (parallel bars). Membranes were clamped at +30 mV throughout the experiments.

multiple-channel membrane. With membranes showing single-channel fluctuations, a ratio of approximately 3.5-4:1 for K $^{+}$ and Na $^{+}$ /Cl $^{-}$ was obtained. The relative permeability of Na $^{+}$ /Ca $^{2+}$ was 12:1. In bi-ionic experiments, K $^{+}$ was slightly favoured (1.2 $\times$ over Na $^{+}$). Thus, FcR $_{\gamma 2b/\gamma 1}$ functions as a monovalent cation-selective channel.

Opening of the FcR-channel may be a signalling event for secretory responses observed in macrophages. Such responses could be due to an initial Na $^{+}$ influx and/or increase in cytoplasmic Ca $^{2+}$ concentration ([Ca $^{2+}$]). Such a rise in [Ca $^{2+}$] could be due to release of internal stores or direct permeation through the occupied receptor (although the permeability of FcR to

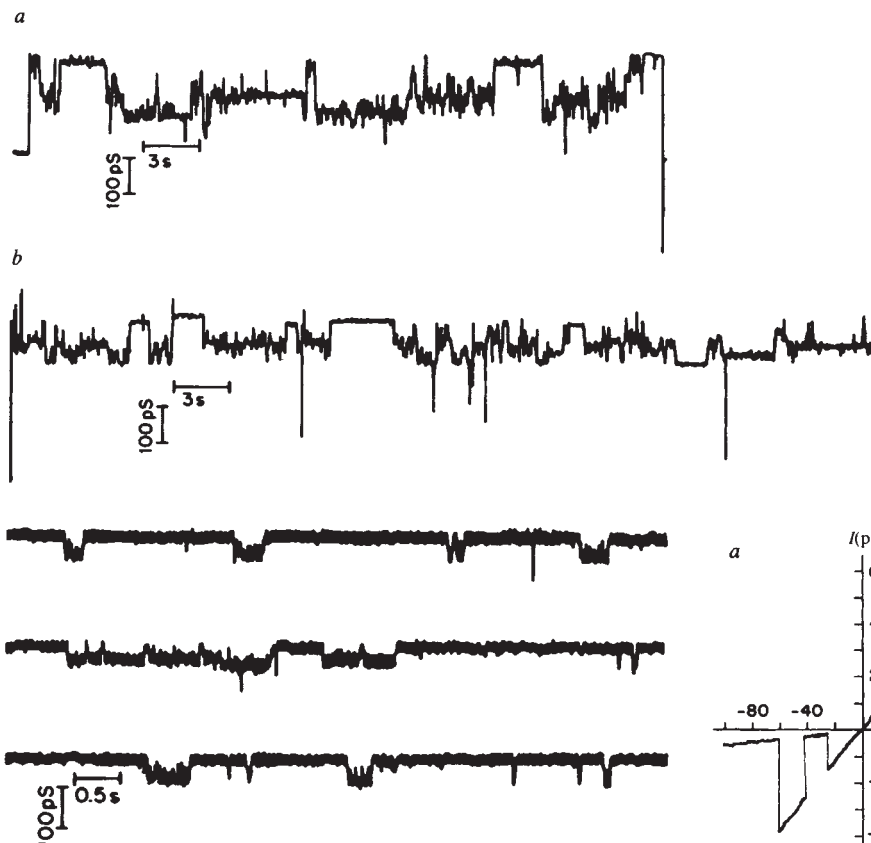


Fig. 4 Single FcR-channel fluctuations activated by 2.4G2 IgG. FcR (1×10^{-11} M) was incorporated into bilayer from *cis* side and challenged with $0.1 \mu\text{M}$ 2.4G2 IgG. The membrane was formed from phosphatidylcholine/choline in 1 M KCl, 3 mM CaCl_2 , 10 mM HEPES, pH 7.4 and clamped at +100 mV. The recording was obtained 5 min after addition of ligand. Downward deflections represent opening of channels. Time resolution was 1 ms.

Ca^{2+} is low: the difference between extracellular $[\text{Ca}^{2+}]$ and $[\text{Ca}^{2+}]_i$ is at least a 1,000-fold²²).

The receptor-ionophore model presented here may relate directly to FcR-mediated phagocytosis of antibody-coated particles, since local rearrangement of the cytoskeleton as a result of localized ion fluxes could lead to the directed membrane movement. A change in local ionic environment could provide the regulatory control for cytoplasmic proteins such as gelsolin^{23,24} which may be actively involved in phagocytosis. Moreover, macrophages display several highly complex rectifying membrane conductances²⁵⁻²⁷, including action potentials²⁸, so our observations support a role for ion fluxes in the normal physiology of these cells.

Ion fluxes through the cell membrane could activate several membrane-associated enzymes. Our data do not preclude the possibility that the receptor itself might display some kind of enzymatic activity following binding. In fact, a phospholipase activity has been reported for this receptor by Suzuki *et al.*²⁹. However the protein they describe has a different M_r and isoelectric point from the FcR described here.

The receptor-ionophore model may be generally applicable to a number of other receptors. Aggregation or clustering of receptors by ligands has also been implied in the functional activation of cells for IgE FcR of mast cells³⁰, epidermal growth factor³¹ and insulin³² receptors. The mouse macrophage FcR may belong to a general class of membrane proteins that are triggered to form ion channels by specific ligands, the prototype of which is the acetylcholine receptor. Properties of the latter have been studied using various techniques, such as patch-clamp³³ and reconstitution into planar bilayers^{34,35}. In contrast to the acetylcholine receptor, which consists of several polypep-

Fig. 3 Discrete current fluctuations following binding of FcR by $0.4 \mu\text{M}$ 2.4G2 IgG. FcR (2×10^{-9} M) was incorporated into bilayers formed in 4 mM MgCl_2 , 3 mM CaCl_2 , 10 mM HEPES buffer, pH 7.4, containing 0.1 M KCl (a) or 0.1 M NaCl (b). Current recordings were initiated by turning the voltage from 0 to +30 mV, 2 min after addition of $0.4 \mu\text{M}$ 2.4G2 IgG to *cis* side. Upward deflections indicate channel openings. Current flowed from *cis* to *trans*. Conductance (vertical bar) was calculated from values of current divided by the membrane voltage.

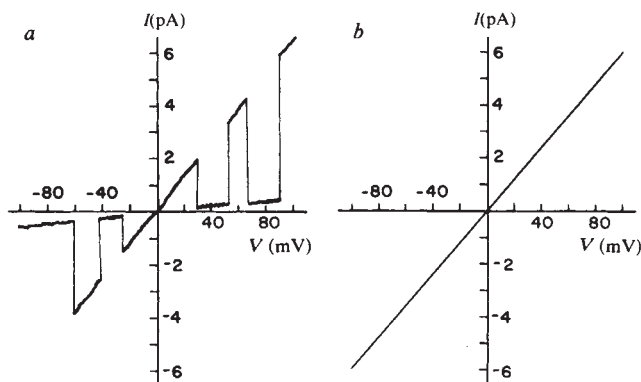


Fig. 5 Open channel I - V relationship obtained by a voltage-clamping ramp. Bilayer formed as for Fig. 4. a, A ramp lasting 700 ms was applied from -100 to +100 mV. Channel fluctuations can be observed during the application of the ramp. b, I - V relationship for single channel reconstructed from a by subtracting the background or voltage-insensitive conductances.

tide chains, however, the mouse IgG FcR contains apparently one single polypeptide⁶. Thus, our observations that a single polypeptide may be modulated by specific ligands to open ion channels presents new opportunities for studying the mechanisms of channel gating. It would be interesting to see whether similar electrical activities are associated with other cells which display complex membrane-associated events.

Preliminary reports of this work have already been presented in abstract form^{36,37}.

This research was supported by NIH grants AI-14603, CA 30198, AI-07012 and Biochemistry Research Support grant 507 RR 07065. J.D.-E.Y. is a recipient of a NIH predoctoral fellowship. T.M.Y. was funded by the National Research Council, Brazil, J.C.U. is an American Cancer Society Research Scholar. We acknowledge helpful discussions with Dr S. Silverstein.

Received 9 June; accepted 25 August 1983.

1. Silverstein, S. C., Steinman, R. M. & Cohn, Z. A. *A. Rev. Biochem.* **46**, 669-722 (1977).
2. Unkeless, J. C., Fleit, H. & Mellman, I. S. *Adv. Immun.* **31**, 247-270 (1981).
3. Nathan, C. F., Murray, H. W. & Cohn, Z. A. *N. Engl. J. Med.* **303**, 622-626 (1980).
4. Cohn, Z. A. *J. Immun.* **121**, 813-816 (1978).
5. Unkeless, J. C. *J. exp. Med.* **150**, 580-596 (1979).
6. Mellman, I. S. & Unkeless, J. C. *J. exp. Med.* **152**, 1048-1069 (1980).
7. Rouzer, C. A., Scott, W. A., Kempe, J. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4279-4282 (1980).
8. Peiris, J. S. M., Gordon, S., Unkeless, J. C. & Porterfield, J. S. *Nature* **289**, 189-191 (1981).
9. Young, J. D.-E., Unkeless, J. C., Kaback, H. R. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1357-1361 (1983).
10. Young, J. D.-E., Unkeless, J. C., Kaback, H. R. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1646-1650 (1983).
11. Racker, E., Vieland, B., O'Neal, S., Alfonzo, M. & Telford, J. *Archs Biochem. Biophys.* **198**, 470-477 (1979).
12. Unkeless, J. C. & Healey, G. A. *J. Immun. Meth.* **56**, 1-11 (1983).
13. Montal, M. & Mueller, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3561-3566 (1972).

14. Young, J. D.-E., Young, T. M., Lu, L. P., Unkeless, J. C. & Cohn, Z. A. *J. exp. Med.* **156**, 1677-1690 (1980).
15. Schindler, H. *FEBS Lett.* **122**, 77-79 (1980).
16. Schindler, H. *Biochim. biophys. Acta* **555**, 316-336 (1979).
17. Young, J. D.-E., Blake, M., Mauro, A. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3831-3835 (1983).
18. Helenius, A., Fries, E. & Kartenbeck, J. *J. Cell Biol.* **75**, 866-880 (1977).
19. Rouzer, C. A. thesis, Rockefeller Univ. (1982).
20. Latorre, R. & Alvarez, O. *Physiol. Rev.* **61**, 77-150 (1981).
21. Finkelstein, A. & Mauro, A. in *Handbook of Physiology - The Nervous System* Vol. 1 (ed. Kandel, E. R.) 161-213 (American Physiological Society, Bethesda, 1977).
22. Rubin, R. P. *Calcium and Cellular Secretion*, 1-276 (Plenum, New York, 1982).
23. Yin, H. L. & Stossel, T. P. *Nature* **281**, 583-586 (1979).
24. Yin, H. L. & Stossel, T. P. *J. biol. Chem.* **255**, 9490-9493 (1980).
25. Gallin, E. K. & Livengood, D. R. *Am. J. Physiol.* **241**, C9-C17 (1981).
26. Gallin, E. K. *Science* **214**, 458-460 (1981).
27. Persechini, P. M., Araujo, E. G. & Oliveira-Castro, G. M. *J. Membrane Biol.* **61**, 81-90 (1981).
28. McCann, F. V., Cole, J. J., Guyre, P. M. & Russell, J. A. G. *Science* **219**, 991-993 (1983).
29. Suzuki, T., Saito-Taki, T., Sadasivan, R. & Nitta, T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 591-595 (1982).
30. Metzger, H. & Ishizaka, T. *Fedn Proc.* **41**, 7-34 (1982).
31. Schechter, Y., Hernaez, L., Schlessinger, J. & Cuatrecasas, P. *Nature* **278**, 835-838 (1979).
32. Kahn, C. R., Baird, K. L., Jarrett, D. B. & Flier, J. S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4209-4213 (1978).
33. Neher, E. & Sakmann, B. *Nature* **260**, 799-802 (1976).
34. Nelson, N., Anholt, R., Lindstrom, J. & Montal, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3057-3061 (1980).
35. Schindler, H. & Quast, U. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3052-3056 (1980).
36. Young, J. D.-E., Young, T. M., Kaback, H. R., Cohn, Z. A. & Unkeless, J. C. *J. gen. Physiol.* **80**, A25 (1982).
37. Young, J. D.-E., Young, T. M., Kaback, H. R., Cohn, Z. A. & Unkeless, J. C. *J. Cell Biol.* **95**, 443a (1982).

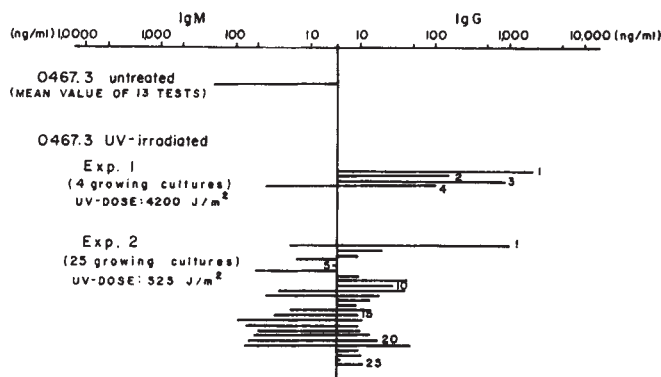


Fig. 1 Effect of UV exposure on immunoglobulin class secretion from 0467.3 cells.

Methods: For UV exposure, 5×10^6 cells were suspended in 2 ml RPMI 1640 medium, layered into a 60×15 mm plastic Petri dish (Falcon 3002), exposed to UV light at a dose rate of $35 \text{ J m}^{-2} \text{ s}^{-2}$. The Petri dishes were placed 100 mm below two Philips TUV 15-W lamps for periods varying between 15 and 120 s. The UV dose rate was determined with a Latarjet dosimeter. After exposure, the cells were transferred to 50-ml plastic flasks (Falcon 3013 F) and cultured at 37°C for 48 h. Dead cells were removed by 1g sedimentation in fetal calf serum (FCS) for 5 h. 3 ml FCS was pipetted into a 15-ml sterile tube. The cells were washed and layered on top of the FCS for sedimentation. The bottom fraction was collected after 5 h and placed in 96-well microtitre plates (Costar, no. 3596), 25,000 cells were seeded per well. The cells were cultured in RPMI 1640 medium containing 10% FCS (Gibco), 100 IU penicillin and $100 \mu\text{g ml}^{-1}$ streptomycin. An enzyme-linked immunosorbent assay was used to measure the quantity of IgM and IgG secreted into the culture medium. 96-well microtitre plates (Dynatech no. 129 A) were coated for 18 h at 25°C with $100 \mu\text{l}$ of rabbit anti-human μ -chain or rabbit anti-human IgG (Fc-specific) at a dilution of 1:500 (IgG fraction of antisera from Dako) in a pH 9.6 carbonate buffer according to Engvall and Perlmann¹¹. $100 \mu\text{l}$ of the supernatants in dilutions were incubated for 1 h at 37°C . Thereafter, $100 \mu\text{l}$ of a horseradish peroxidase conjugated anti-human μ or anti-human IgG were added in a dilution of 1:400 or 1:600, respectively. After each incubation, the plates were washed in phosphate buffer, pH 7.4, containing 0.05% Tween 20. Standard samples with known amounts of IgM or IgG were from Behringwerke AG and included concentrations of $500 \mu\text{g ml}^{-1}$ to 0.2 ng ml^{-1} of immunoglobulin in each plate. $200 \mu\text{l}$ of the substrate solution [2 mM ABTS (Boehringer Mannheim), 2.5 mM H_2O_2 in a 0.1 M Na-acetate/0.05 M NaH_2PO_4 buffer] was added for 30 min at 37°C . The reaction was stopped by the addition of $25 \mu\text{l}$ 10% SDS. The colour intensity was measured at 405 nm in a Titertek Multiskan multichannel spectrophotometer.

technique. In expt 1, IgG was found in the supernatant of 4 of 4 growing cultures and in expt 2, in 21 of 25 growing cultures. Subsequent analyses of the cells obtained in expt 2 showed IgG secretion in all growing cultures. One of the 4 growing cultures in expt 1 and 13 of the 25 growing cultures in expt 2 showed some IgM secretion in addition to IgG secretion. The four UV-exposed cultures from the expt 1 that switched to IgG secretion have now been maintained in culture for 5 months. All produce IgG and the single double producer continues to make IgM λ as well. The cells from expt 2 have been in culture for 8 weeks. Further experiments performed at various UV doses have confirmed that the IgG induction is a repeatable event, and that stable IgG secreting lines can be established by subcloning. Lymphoblastoid cell lines are generally relatively stable with regard to immunoglobulin class and the quantity of immunoglobulin produced², although shifts can occur³. The detailed mechanisms and regulation of the constant region heavy-chain (C_H) gene switching are unknown, although it is believed that the class switch is due to the activation of a C_H -class specific switch recombinase at a certain stage of the B-cell differentiation⁴. T lymphocytes or their products appear to induce the switch, or select specific C_H isotypes from populations of normal B cells⁵.

UV light-induced immunoglobulin heavy-chain class switch in a human lymphoblastoid cell line

Anders Rosén & George Klein

Department of Tumor Biology, Karolinska Institutet, S-104 01 Stockholm, Sweden

During attempts to select nonsecretory variants from 0467.3, an Epstein-Barr virus-transformed human lymphoblastoid cell line¹ that secretes small amounts of IgM λ , we exposed the cells to UV light. Cells that survived the irradiation were subcultured and their supernatants were screened for immunoglobulin production by an enzyme-linked immunosorbent assay (ELISA). Although stable nonsecretory variants were not isolated, we report here that an immunoglobulin class switch occurred in the UV-treated cell population. All survivors were found to produce large quantities of IgG λ . Some cell cultures also produced the original IgM λ . The UV-light-induced class switch was regularly reproducible with this target cell line.

The 0467.3 line is a subclone¹ of an hypoxanthine-guanine phosphoribosyl transferase-deficient Epstein-Barr virus-transformed B-cell line (GM 0467) derived from the peripheral blood of a patient with infectious mononucleosis. The line has been cloned several times. Between 1 and 3% of the cells express surface IgM, and 13-21% contain intracellular μ, λ , in parallel with some plasmacytoid differentiation. Three-day-old supernatants from 0467.3 cells in logarithmic growth were found to contain $150 \text{ ng IgM per ml}$ (mean value of 13 tests), as measured by ELISA. In UV exposure experiments, cells were treated with a UV dose of $4,200 \text{ J m}^{-2}$ (expt 1) or 525 J m^{-2} (expt 2). Five days after the UV exposure, the cells were examined microscopically. Approximately 1% of the cells were alive. Three weeks after the initiation of the cultures, growing clones were seen. The cells were fed every second day and the 2-day-old supernatants were collected and tested for IgG and IgM by ELISA.

Figure 1 shows the IgG and IgM levels in the supernatants collected from wells that contained growing cells. While the supernatants of the untreated 0467.3 line in 13 independent tests showed IgM but no detectable IgG secretion, all UV-treated growing cultures produced IgG. No IgG secreting variants of the 0467.3 line could be detected to arise spontaneously, as analysed by screening by ELISA or at the single-cell level by immunofluorescence and the protein A-plaque

Ultraviolet exposure leads to the formation of pyrimidine dimers in the DNA. A complex equipment of repair enzymes can correct this damage and maintain the integrity of the DNA. It is conceivable that some of these enzymes are instrumental in causing the appropriate C_H deletion expressed as the IgM to IgG switch. Alternatively, the enzymes responsible for the hypermutability of the variable region heavy-chain gene (V_H) may be involved. It is known that both localized base substitutions (specific mutations) and rearrangements occur in the V_H region⁶. While the mechanism for the hypermutation is not known, it is noteworthy that IgG, but not IgM, V_H domains contain new sequences due to mutations that are not encoded in the germ lines⁷. This implies that the enzyme system responsible for the mutations, envisaged by some as an error-prone repair system^{8,9}, act in conjunction with the heavy-chain switch from IgM to IgG. These enzymes may belong to the same category as the enzymes concerned with the repair of UV-damaged DNA. As a third alternative, DNA damage at a critical

stage of B-cell development may interfere with the natural differentiation process and induce an immunoglobulin heavy-chain class switch.

Superficially at least, the $\mu \rightarrow \gamma$ switch in all UV-exposed 0467.3 cultures is reminiscent of the UV-induced production of phage in lysogenic bacterial cultures¹⁰. This is particularly suggested by the regularity of the event, in contrast to mutation-like phenomena.

In conclusion, our findings suggest that UV illumination may be used for the regular induction of an IgH class switch in human lymphoblastoid cell lines.

This investigation was supported by PHS grant 2 RO1 CA 25250-03 awarded by the National Cancer Institute, DHHS, the Swedish Cancer Society and by funds from the Karolinska Institute, Swedish Board of Technical Development and the Department of Agriculture. We thank Ms Gunilla Kjellström and Ms Daniele Schmidt for technical help, and Dr Jack Spira for helpful discussion.

Received 23 May; accepted 5 September 1983.

- Chiorazzi, N., Wasserman, R. L. & Kunkel, H. G. *J. exp. Med.* **156**, 930-935 (1982).
- Rosén, A., Clements, G., Klein, G. & Zeuthen, J. *Cell* **11**, 139-147 (1977).
- Saiki, O., Kishimoto, T., Kuritani, T., Muraguchi, A. & Yamamura, Y. *J. Immun.* **124**, 2609-2614 (1980).
- Sher, I., Ahmed, A., Strong, D. M., Steinberg, A. D. & Paul, W. E. *J. exp. Med.* **141**, 788-803 (1975).

- Isakson, P., Puré, C., Vitetta, E. S. & Krammer, P. H. *J. exp. Med.* **155**, 734-748 (1982).
- Kim, S., Davis, M., Sinn, E., Patten, P. & Hood, L. *Cell* **27**, 573-581 (1981).
- Gearhart, P. J., Johnson, N. D., Douglas, R. & Hood, L. *Nature* **291**, 29-34 (1981).
- Selsing, E. & Storb, U. *Cell* **25**, 47-58 (1981).
- Brenner, S. & Milstein, C. *Nature* **211**, 242-243 (1966).
- Lwoff, A., Simionovitch, L., & Kjeldgaard, N. *Ann. Inst. Pasteur* **79**, 815-858 (1953).
- Engvall, E. & Perlmann, P. *Immunochemistry* **8**, 871-874 (1971).

Functional expression of a transfected murine class II MHC gene

Ronald N. Germain & Michael A. Norcross

Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205, USA

David H. Margulies*

Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20205, USA

The activation of T helper lymphocytes involves the recognition of class II major histocompatibility complex antigens, which are dimeric glycoproteins (of subunit composition $A_\alpha A_\beta$ or $E_\alpha E_\beta$) expressed on the surfaces of macrophages and B lymphocytes¹⁻⁵. One approach to understanding the relationship between the structure of these antigens and their functions in the immune response^{6,7} is to clone the genes that encode them⁸⁻¹³, to obtain functional expression of the cloned genes transfected into an appropriate cell line, and then to see how those functions are affected in variant genes generated *in vitro*. We report here the expression in Ia^d -bearing B cells of an A_β^k gene, which confers on the transfected cells the capacity for both allostimulation and antigen-dependent activation of an $I-A^k$ -restricted T-cell clone.

The primary target cells in most gene transfer experiments reported so far have been fibroblasts¹⁴. Although class II major histocompatibility complex (MHC) genes have been successfully introduced into, and expressed by such cells (ref. 15 and our own unpublished data), in contrast to experiments with class I genes¹⁶⁻¹⁸, fibroblasts are not the most suitable cell type in which to study the function of class II MHC molecules. Other laboratories have reported transfection experiments using an appropriately differentiated cell type as the recipient of the transfected gene, for example myeloma cells¹⁹ and pre-B cells²⁰ have been transfected with immunoglobulin genes. We have adopted this approach, and chosen cloned B-cell tumour lines,

which constitutively express class II MHC (Ia) molecules and are capable of antigen presentation to T cells^{21,22}, as transfection recipients of a murine Ia gene.

Two plasmids were constructed (see Fig. 1), each containing a complete A_β^k gene¹³ together with the prokaryotic xanthine-guanine phosphoribosyl transferase (*Ecogpt*) gene²³. The *Ecogpt* gene serves as a dominant marker, so that stable transfectants can be selected in a medium containing inhibitors of the normal purine synthetic pathway. Plasmid DNA was introduced into the antigen-presenting, Ia^+ BALB/c ($I-A^d$, $I-E^d$) derived, hypoxanthine-guanosine phosphoribosyl transferase-negative B-cell lymphoma M12.4.1²¹, using a modification of the spheroplast fusion method of Sandri-Goldin *et al.*²⁴. Transformants containing functional *Ecogpt* genes were selected for in mycophenolic acid-xanthine-containing medium. After 1-3 weeks, several clones of growing cells were seen. These clones were expanded and studied for integration and expression of A_β^k genes.

To demonstrate the transfer of the A_β^k gene into the selected clones, high molecular weight DNA was prepared from two such putative transformant clones, T70.3.1 and T70.4.1, which were transfected with plasmids pI- A_β^k -gpt-1 and pI- A_β^k -gpt-2 respectively. This DNA, in parallel with DNA from BALB/c ($I-A^d$) cells and from the T-cell hybridoma 2B4 (genotypically $I-A^k$), was digested with *Bam*HI restriction endonuclease, and Southern blotted²⁵ after agarose gel electrophoresis. The blot was probed with a 150 base pair (bp) *Pst*I fragment of cDNA clone pI- A_β^d -1, which hybridizes to a 2.1 kb *Bam*HI fragment of the A_β^k allele, and a 9.5 kb *Bam*HI fragment of the A_β^d allele. As shown in Fig. 2, both T70.3.1 and T70.4.1 contain, in addition to the M12.4.1-derived endogenous A_β^d 9.5 kb fragment, a major new fragment of 2.1 kb, co-migrating with the authentic A_β^k fragment from 2B4. On the basis of relative intensities of the 2.1 kb A_β^k band and the endogenous 9.5 kb A_β^d band, T70.3.1 appears to have several copies of A_β^k , while T70.4.1 has only one or two copies of the transfected gene.

These experiments indicated the successful introduction of the A_β^k recombinant plasmids into M12.4.1 DNA. As an initial test for the expression of the transfected gene, the cells were stained with a variety of monoclonal anti- Ia antibodies and fluorescent rabbit anti-mouse IgG, then analysed by flow microfluorimetry. Figure 3 shows the staining profiles of T70.3.1 and T70.4.1, compared with those of M12.4.1, TA-3 (an M12.4.1-derived $Ia^k \times Ia^d$ hybridoma)²⁶, and CH1 (an Ia^{k+} B-cell tumour)²⁷ cells. The $I-A^d$ -specific antibody MKD6²⁸ stains all

* Present address: Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205, USA.

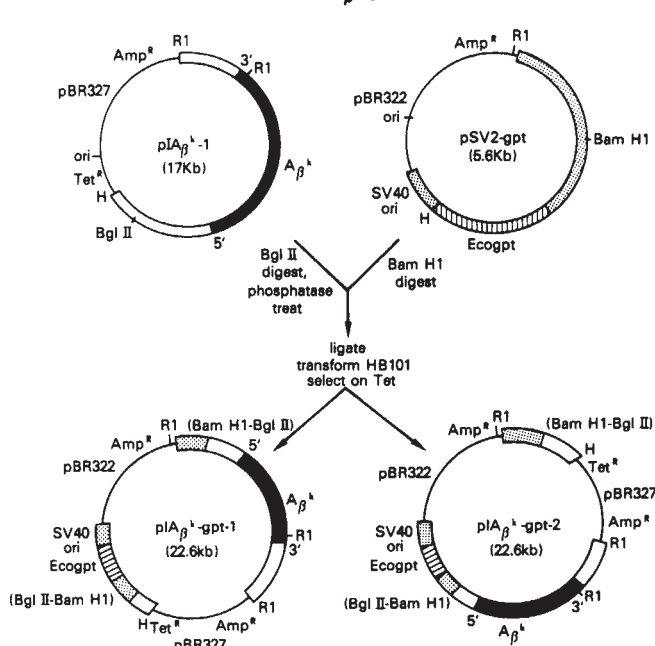
Construction of A_{β}^k gpt Vectors

Fig. 1 Construction of recombinant plasmids containing the A_{β}^k and *Ecogpt* genes. An 800 bp *Cfo*I fragment of a cDNA clone¹² containing a 468 bp insert corresponding to the 3' untranslated, intracytoplasmic and transmembrane regions of A_{β}^d mRNA was nick-translated. The resultant ³²P-labelled probe was used to identify hybridizing genomic fragments in a partial *Hae*III library of A/J (H-2^a, I-A^k) liver DNA cloned via *Eco*RI linkers in the λ phage vector Charon 4A. One hybridizing clone, A2.3.4.1, was found to contain a 21 kb insert including the entire A_{β}^k gene plus flanking DNA. A 9.6 kb *Hind*III-*Eco*RI fragment and a 3.7 kb *Eco*RI fragment, containing the 5' and 3' portions of the A_{β}^k gene, respectively, were subcloned into pBR327. The 3.7 kb 3' *Eco*RI fragment was recovered from this plasmid and ligated to *Eco*RI linearized plasmid DNA of the 5' subclone. The ligated DNA was used to transform *Escherichia coli* MC1061. Miniprep DNA from colonies growing on tetracycline-containing LB agar plates was used to determine the orientation of the inserted 3' fragment. One such construct with a complete A_{β}^k gene was isolated and termed pI- A_{β}^k -1. This plasmid was linearized by digestion with *Bgl*II, which recognizes a unique site several kilobases 5' of the A_{β}^k gene proper. This linearized DNA was treated with alkaline phosphatase to prevent self-ligation during the subsequent steps. The plasmid vector pSV2gpt²³ containing the *Ecogpt* gene was linearized with *Bam*HI. These two linear DNAs were ligated to each other, and used to transform competent *E. coli* HB101. The transformed bacteria were plated on LB agar containing tetracycline. This prevented the growth of bacteria containing self-ligated pSV2gpt DNA, which only contains the gene for ampicillin resistance. Several resistant colonies were picked and miniprep DNA restriction mapped to determine if the complete A_{β}^k gene was present, and its orientation with respect to the pSV2gpt vector sequences. Two plasmids with A_{β}^k in opposite orientations (pI- A_{β}^k -gpt-1 and pI- A_{β}^k -gpt-2) were all selected for use in these experiments. All procedures involved in the construction of these recombinant plasmids were performed essentially according to Maniatis *et al.*³⁷. Abbreviations: Amp^r, ampicillin resistance gene; Tet^r, tetracycline resistance gene; RI, *Eco*RI site; BglII, *Bgl*II site; H, *Hind*III site; Bam HI, *Bam*HI site; ori, origin of replication; *Ecogpt*, *E. coli* xanthine-guanine phosphoribosyl transferase gene. Black regions represent the A_{β}^k gene; stippled regions represent SV40 sequences; lined regions represent the *Ecogpt* gene; white regions represent DNA flanking A_{β}^k ; and thin lines represent plasmid sequences. The various plasmids are not drawn to scale, but the relative sizes of the different regions in each construct are approximately correct.

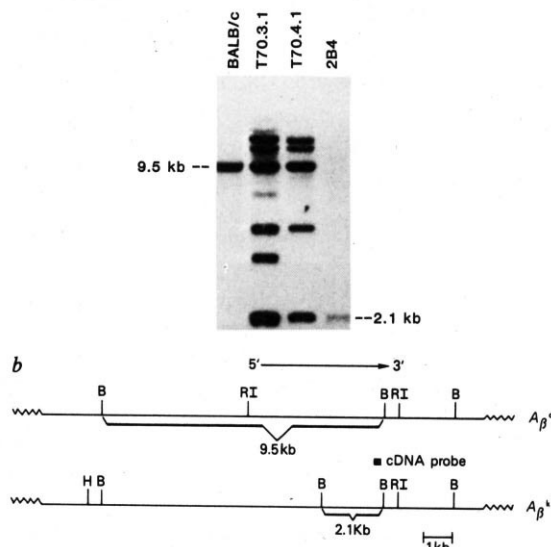


Fig. 2 Identification of I- A_{β}^k genes in transformed B cells. *a*, The figure shows a Southern blot of DNA from cells transfected with the cloned A_{β}^k gene (lanes T70.3.1 and T70.4.1), and from BALB/c and T-cell hybridoma 2B4 cells (lanes BALB/c and 2B4 respectively) hybridized to a ³²P-labelled A_{β} gene probe. *b*, Restriction maps of the A_{β}^d and A_{β}^k genes are shown with respect to the 150 bp cDNA fragment used as probe. Endonuclease cleavage sites: H, *Hind*III; B, *Bam*HI; RI, *Eco*RI.

Methods: I- A_{β}^k recombinants pI- A_{β}^k -gpt-1 and pI- A_{β}^k -gpt-2 as described in Fig. 1 legend were introduced into M12.4.1 cells by spheroplast fusion using the method of Sandri-Goldin *et al.*²⁴ as described by Oi *et al.*¹⁹. Spheroplasts of *E. coli* strain HB101 harbouring the indicated plasmid were prepared as described previously¹⁹. 15 ml were added to 1×10^7 M12.4.1 cells in 9 ml of medium and centrifuged at 500g for 5 min. Following careful aspiration of the supernatant, cells were resuspended for 3 min in 2 ml Dulbecco's modified Eagle medium (DME) containing 50% (v/v) polyethylene glycol, 50 mM Tris-HCl (pH 8.0). The polyethylene glycol/spheroplast mixture was promptly diluted out by the addition of 8 ml of serum-free DME. Cells were centrifuged again, resuspended in DME containing 20% fetal bovine serum, supplemented with 2 mM glutamine, non-essential amino acids, 50 μ g ml⁻¹ gentamicin, 10^{-5} M 2-mercapto-ethanol, 0.5 mM sodium pyruvate, 1 mM oxaloacetate, 0.2 U ml⁻¹ insulin and 10% NCTC 109 (DME-complete)³⁸. Cells were resuspended to a density of 1×10^6 per ml and incubated for 24 h. At this time cells were fed with an equal volume of the same medium, and selection was imposed by the addition of mycophenolic acid, hypoxanthine and xanthine to 6 μ g ml⁻¹, 15 μ g ml⁻¹ and 250 μ g ml⁻¹ respectively. Cells were distributed in 0.1 ml selective medium to 96 well tissue culture plates and maintained at 37°C in a 10% CO₂ atmosphere. 100 μ l of selective medium was added once weekly, and after 2-3 weeks, several colonies were identified and grown to mass culture for further analysis. T70.3.1 and T70.4.1 were independent clonal isolates resulting from transfection with pI- A_{β}^k -gpt-1 and pI- A_{β}^k -gpt-2 respectively. To analyse the transformant cells for the presence of the transfected genes, high molecular weight DNA was prepared by a modification of the method of Blin and Stafford³⁹. This was subjected to cleavage with the restriction endonuclease *Bam*HI, and 10 μ g were fractionated by electrophoresis in 0.8% agarose gel in a buffer containing 40 mM Tris, 20 mM Na acetate, and 40 mM EDTA (pH 7.2). DNA was denatured *in situ*, transferred to nitrocellulose paper²⁵, and hybridized (at 40°C for 24 h) to the ³²P-labelled 150 bp *Pst*I fragment derived from the pI- A_{β}^d -1 cDNA clone¹², which had been nick-translated to a specific activity of 2.0×10^8 d.p.m. per μ g. The blot was washed in $2 \times$ SSC (0.300 M NaCl, 0.30 M Na citrate), 0.1% SDS at 25°C, then in $0.1 \times$ SSC (0.015 M NaCl, 0.0015 M Na citrate), 0.1% SDS at 50°C, dried and subjected to autoradiography with XAR-5 (Kodak) film and intensifying screens for 35 h at -80°C. Nonparental hybridizing fragments may represent minor pBR322 contamination of the A_{β} probe or may result from recombination between plasmids or between plasmid and cellular DNA during the process of gene transfer⁴⁰. Similar experiments using M12.4.1 DNA gave identical results to the BALB/c DNA in this blot (not shown). Molecular size (kb) is interpolated from that of co-migrated bacteriophage markers.

Table 1 T-cell responses to A_β^k transfectants

Responder cells	APC	No antibody	10.2.16	M5/114	Both	
Expt 1*						
Whole lymph node	None	6,864 $\pm$ 1,162	9,837 $\pm$ 979	2,185 $\pm$ 280	—	
	M12.4.1	24,682 $\pm$ 1,858	31,224 $\pm$ 2,643 (–26)	7,625 $\pm$ 659 (69)	8,855 $\pm$ 864 (64)	
	T70.3.1	73,567 $\pm$ 4,180	22,050 $\pm$ 336 (70)	41,073 $\pm$ 2,940 (44)	6,058 $\pm$ 292 (92)	
	T70.4.1	39,203 $\pm$ 1,504	16,165 $\pm$ 2,172 (59)	14,919 $\pm$ 858 (62)	4,766 $\pm$ 3,370 (88)	
	TA3	112,433 $\pm$ 22,217	68,746 $\pm$ 300 (39)	31,061 $\pm$ 845 (72)	29,327 $\pm$ 1,637 (74)	
Expt 2†						
Lymph node	—	372 $\pm$ 63	484 $\pm$ 104	456 $\pm$ 50	408 $\pm$ 73	
T cells	M12.4.1	17,926 $\pm$ 642	18,888 $\pm$ 4,730 (–4)	2,253 $\pm$ 353 (87)	2,047 $\pm$ 451 (89)	
	T70.3.1	66,608 $\pm$ 4,408	12,766 $\pm$ 2,412 (81)	23,978 $\pm$ 5,032 (64)	1,338 $\pm$ 333 (98)	
	TA3	313,730 $\pm$ 19,944	258,016 $\pm$ 2,029 (18)	230,433 $\pm$ 6,074 (27)	30,377 $\pm$ 2,961 (90)	
Expt 3‡						
	APC	No addition	GAT	GAT+10.2.16	GAT+M5/114	GAT+both
GAT-specific	—	356 $\pm$ 2	190 $\pm$ 41	154 $\pm$ 63	114 $\pm$ 14	180 $\pm$ 30
T cell clone	M12.4.1	389 $\pm$ 48	430 $\pm$ 107	377 $\pm$ 58	255 $\pm$ 57	304 $\pm$ 8
	T70.3.1	1,229 $\pm$ 333	13,177 $\pm$ 2,537	1,239 $\pm$ 186 (91)	20,001 $\pm$ 4,668 (–52)	799 $\pm$ 132 (94)
	TA3	2,050 $\pm$ 530	14,288 $\pm$ 479	891 $\pm$ 172 (94)	21,569 $\pm$ 1,956 (–51)	452 $\pm$ 123 (97)

* A single cell suspension of BALB/c (H-2^d, I-A^d) lymph node cells was prepared in EHAA: RPMI-1640, 1:1, containing 10% fetal calf serum, 2 mM L-glutamine, 10 mM HEPES, 5×10^{-5} M 2-mercaptoethanol, and penicillin-streptomycin (ER medium). 2.5×10^5 viable lymph node cells were cultured in 200 μ l total volume in the wells of flat-bottomed microtitration dishes, with or without the addition of 10^4 10,000R γ -irradiated tumour cells (APC), in the presence or absence of 25% (v/v) 10.2.16 hybridoma supernatant, 12.5% M5/114 hybridoma supernatant, or both supernatants together. After 3 days of culture in 5% CO₂ at 37°C, each culture well received 1 μ Ci ³H-thymidine (6.9 Ci mmol^{–1}, NEN). 24 h later the cells were collected onto glass fibre filters using an automated sample collector, and the samples prepared for liquid scintillation counting. Data are expressed as total mean c.p.m. $\pm$ s.e. for triplicate cultures. Numbers in parentheses represent % inhibition of response by monoclonal antibodies, calculated as % inhibition = (c.p.m. for no antibody – c.p.m. with antibody/c.p.m. for no antibody) $\times$ 100%.

† BALB/c lymph node cells prepared as for expt 1 were passed through a nylon wool column and treated with M5/114 plus complement as previously described²⁸. The resultant purified T cells showed virtually no residual response to the mitogen concanavalin A, but responded normally on addition of phorbol myristate acetate plus mitogen (data not shown). 2.5×10^5 of these purified T cells were used per culture as responders according to the experimental design detailed for expt 1 above.

‡ 1.5×10^4 cells of the B10.A derived, GAT-specific, I-A^k restricted T-cell clone 11.4 were incubated as described above, but with or without the addition of 100 μ g ml^{–1} of GAT, and only 10^3 instead of 10^4 APC. These cultures were pulsed with ³H-thymidine 48 h after initiation. Data are expressed as above.

cells, except for CH1, as expected. The A_β^k -specific antibody 10.2.16²⁹ fails to stain the A_β^d -bearing M12.4.1 parent cells, but reacts strongly with the A_β^k transfectants, as well as the I-A^{k+} control cells, TA-3 and CH1. The A_α^k -specific monoclonal antibody 39J³⁰ stains the I-A^k cells, but does not stain either the M12.4.1 or the two transfectants. These results rule out the possibility that clones T70.3.1 and T70.4.1 are contaminants derived from TA-3 or other Ia^k $\times$ Ia^d hybridomas maintained in this laboratory, and confirm the A_β^k specificity of the monoclonal antibody 39J. The heterogeneity in level of Ia expression seen in the fluorescence profiles was similar for the parent cells and transfectants, and so is not the result of the gene transfer procedure. Based on staining with the 10.2.16 antibody (and also with the A_β^k -specific antibody 40N³⁰, data not shown), the level of A_β^k expression on the transfectants is similar to that of several I-A^{k+} B-cell tumours, and also tends to be greater on the T70.3.1 than T70.4.1 cells. This may reflect the additional number of A_β^k genes in the T70.3.1 line. We presume that the observed surface expression of A_β^k involves an A_β^k : A_α^d heterodimer, as Ia β chains are not expressed on the cell surface in the absence of the appropriate α chains³¹.

Having demonstrated the cell-surface expression of A_β^k in a form reactive with highly specific monoclonal antibodies, we analysed the transfectants for their ability to be recognized by T lymphocytes via this new cell-surface determinant. Initial experiments analysed the clones for their expression of allogeneic Ia determinants as detected by a mixed lymphocyte reaction (MLR). As shown in Table 1, expt 1, both T70.3.1 and T70.4.1 were able to stimulate BALB/c (H-2^d) lymph node cells, syngeneic with the parental M12.4.1 tumour, to undergo extensive proliferation. However, as previously documented, the M12.4.1 itself stimulates a strong autologous MLR with BALB/c responders²². To distinguish the autologous MLR from allostimulation due to A_β^k expression, blocking studies using monoclonal anti-Ia antibodies were performed. Whilst anti-I-A^d, I-E^d (M5/114)^{32,33}, but not anti-I-A^k (10.2.16) antibodies blocked the autologous MLR stimulated by M12.4.1, anti- A_β^k

antibodies had a marked inhibitory effect on the proliferation induced by the transfectants. This documented the role of the transferred and expressed A_β^k gene in stimulating the lymphocytes to proliferate. To rule out that A_β^k shed from the membranes of the transfectants was being presented by I-A^{d+} macrophages in the lymph node preparations³⁴, as well as to identify the responding cell population, T lymphocytes were purified by nylon wool filtration and treatment with anti-Ia^d plus complement (expt 2). These purified T cells (which, due to the absence of Ia⁺ antigen presenting cells, lacked responsiveness to concanavalin A)³⁵, were strongly stimulated by the A_β^k transfectants, and this stimulation was inhibited by the anti- A_β^k monoclonal antibody. Thus, the expression of A_β^k alone, in the absence of A_α^k , generates determinants recognized as allogenic by unprimed T lymphocytes.

To determine if, in addition to allogeneic MLR stimulation, A_β^k in the absence of A_α^k could serve as a 'restricting element' for I-A^k plus antigen-specific T lymphocytes, a B10.A (H-2^a) T-cell clone maintained in IL-2 in the absence of antigen presenting cells was used (J. Ashwell, personal communication). This I-A^k restricted, cloned T cell fails to proliferate in response to I-A^{d+} M12.4.1 cells plus the appropriate antigen L-glutamic acid⁶⁰-L-alanine³⁰-L-tyrosine¹⁰ (GAT). However, this same T-cell clone responds equally well to T70.3.1 or control TA-3 cells in the presence of antigen. The proliferation can be blocked by anti- A_β^k , but not anti-I-A^d, antibodies.

The results presented above indicate that class II MHC genes can be introduced stably (persistent for over 100 generations in culture) into target cells physiologically capable of T-cell stimulation, and further, that the products of such a transfected gene are expressed on the cell surface in a form recognizable by both specific antibodies and T lymphocytes. Expression of A_β^k by a cell already producing A_β^d , A_α^d , E_β^d , E_α^d , but not A_α^k , results in acquisition of both allodeterminants and at least certain I-A^k-specific restriction elements required for antigen specific T-cell activation. Future experiments will ascertain whether the introduction of A_α^k into A_β^k transfectants results in the develop-

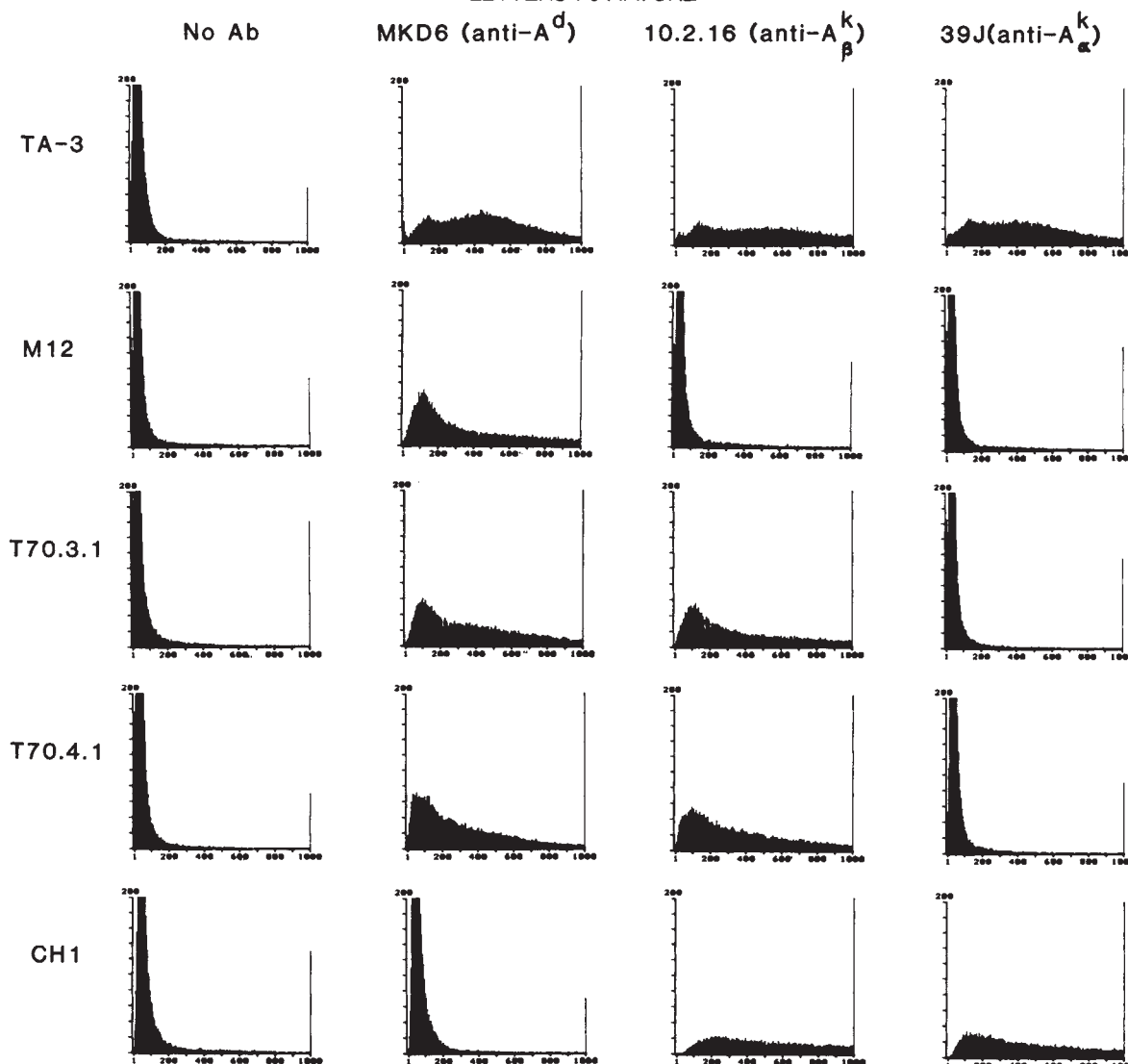


Fig. 3 Surface $I-A^k_\beta$ expression by transformed cells. The figure shows the results of flow microfluorimetry analyses of the cell lines TA3 (a B-cell hybridoma derived by the fusion of BALB/c $\times$ A/J F_1 ($H-2^d \times H-2^k$) B cells to M12.4.1 cells), M12 (M12.4.1, a BALB/c B cell lymphoma²⁶), T70.3.1, T70.4.1 and CH1 (a $H-2^k$ lymphoma) with the indicated antibodies.

Methods: 1×10^6 of each of the cell lines were washed with cold phosphate-buffered saline containing 0.5% fetal bovine serum, 0.5% bovine serum albumin and 1 mM NaN_3 (PBSA), then incubated with either DME 10% fetal bovine serum (no antibody) or with the indicated monoclonal antibody (MKD6 as the culture supernatant, 10.2.16 and 39J as the purified protein) in antibody excess for 30 min at 4 °C. Cells were washed twice again, then incubated for an additional 30 min with a fluoresceinated goat $F(ab')_2$ anti-mouse IgG, washed twice again, resuspended in 1 ml of PBSA and analysed by flow microfluorimetry using an Ortho cytofluorograph. The abscissa plots the fluorescence intensity as a linear function of the number of cells. (A total of 2.5×10^4 cells was analysed for each tracing.) Assignments for the specificity of the indicated monoclonal antibodies are based on references 28–30 and L. Glimcher (personal communication).

ment of additional allodeterminants or restriction elements not possessed by A^k_β in its presumed expression with A^d_α . The approach described here, applied to *in vitro* variants of class II genes created by exon shuffling³⁶ or site-directed mutagenesis, should permit the fine mapping of the structural features of Ia molecules critical to their immunological function.

We thank Kathy Nickerson, Jenny Tou, J. Ashwell, S. Morrison, Howard Holden, Michael Walker, E. Hanna, R. Hodes, Shirley Starnes, R. Schwartz and W. Paul for their help.

Received 18 August; accepted 15 September 1983.

- Schwartz, R. in *Ia Antigens* Vol. 1 (eds Ferrone, S. & David, C. S.) 161–218 (CRC Press, Boca Raton, 1982).
- Nagy, Z. A., Baxevanis, C. N., Ishii, N. & Klein, J. *Immun. Rev.* **60**, 59–83 (1981).
- Ziegler, K. & Unanue, E. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 175–178 (1981).
- Germain, R. N. *J. Immun.* **127**, 1964–1966 (1981).
- Chestnut, R. W., Colon, S. M. & Grey, H. *J. Immun.* **129**, 2383–2388 (1982).
- Benacerraf, B. & Germain, R. N. *Immun. Rev.* **38**, 70–119 (1978).
- Shreffler, D. C. & David, C. S. *Adv. Immun.* **20**, 125–195 (1975).
- Steinmetz, M. *et al. Nature* **300**, 35–43 (1982).
- McNicholas, J., Steinmetz, M., Hunkapiller, T., Jones, P. & Hood, L. *Science* **218**, 1229–1232 (1982).

- Mathis, D. J., Benoist, C. O., Williams, V. E., Kanter, M. R. & McDevitt, H. O. *Cell* **32**, 745–754 (1983).
- Benoist, C. O., Mathis, D. J., Kanter, M. R., Williams, V. E. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **80**, 534–538 (1983).
- Robinson, R., Germain, R. N., McKean, D., Mescher, M. & Seidman, J. G. *J. Immun.* **131**, 2025–2031 (1983).
- Choi, E., McIntyre, K., Germain, R. N. & Seidman, J. G. *Science* **221**, 283–286 (1983).
- Scangos, G. & Ruddle, F. H. *Gene* **4**, 1–10 (1981).
- Rabourdin-Combe, L. & Mach, B. *Nature* **303**, 670–674 (1983).
- Margulies, D. H. *et al. J. Immun.* **130**, 463–470 (1983).
- Reiss, C. S., Evans, G. A., Margulies, D. H., Seidman, J. G. & Burakoff, S. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2709–2712 (1983).
- Forman, J., Goodenow, R. S., Hood, L. & Ciavarella, R. *J. exp. Med.* **157**, 1261–1272 (1983).
- Oi, V. T., Morrison, S. L., Herzenberg, C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **80**, 825–829 (1983).
- Rice, D. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7862–7865 (1982).
- Hamano, T., Kim, K. J., Leiserson, W. M. & Asofsky, R. *J. Immunol.* **129**, 1403–1406 (1982).
- Glimcher, L. H. *et al. Nature* **298**, 283–284 (1982).
- Mulligan, R. C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2072–2076 (1981).
- Sandri-Goldin, R. M., Goldin, A. L., Levine, M. & Glorioso, J. C. *Molec. cell Biol.* **1**, 743–752 (1981).
- Southern, E. J. *molec. Biol.* **98**, 503–517 (1975).
- Glimcher, L. H., Hamano, T., Asofsky, R. *et al. J. Immun.* **130**, 2287–2294 (1983).
- Lynes, M. A., Lanier, L., Babcock, G. F., Wettstein, P. J. & Haughton, G. *J. Immun.* **121**, 2352–2357 (1978).
- Nepom, J. T., Benacerraf, B. & Germain, R. N. *J. Immun.* **127**, 31–34 (1981).

29. Oi, V. T., Jones, P. P., Goding, J. W., Herzenberg, L. A. & Herzenberg, L. A. *Curr. Topics microbiol. Immun.* **81**, 115-129 (1978).
30. Pierres, M., Devaux, C., Dosseto, M. & Marchetto, S. *Immunogenetics* **14**, 481-495 (1981).
31. Murphy, D. B., Jones, P. P., Loken, M. R. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5404-5408 (1980).
32. Bhattacharya, A., Dorf, M. E. & Springer, T. A. *J. Immun.* **127**, 2488-2495 (1981).
33. Germain, R. N., Bhattacharya, A., Dorf, M. E. & Springer, T. A. *J. Immun.* **128**, 1409-1413 (1982).
34. Weinberger, O., Germain, R. N., Springer, T. & Burakoff, S. J. *J. Immun.* **129**, 694-697 (1982).
35. Habu, S. & Raff, M. C. *Eur. J. Immun.* **7**, 451-457 (1977).
36. Evans, G. A., Margulies, D. H., Shykind, B., Seidman, J. G. & Ozato, K. *Nature* **300**, 755-757 (1982).
37. Maniatis, T., Fritsch, E. T. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
38. Ozato, K., Mayer, N. & Sachs, D. H. *J. Immun.* **124**, 533-540 (1980).
39. Blin, N. & Stafford, D. W. *Nucleic Acids Res.* **3**, 2303-2308 (1976).
40. Anderson, R. A., Krakauer, T. & Camerini-Otero, R. D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2748-2752 (1982).

Amplification and expression of the *c-myc* oncogene in human lung cancer cell lines

Cameron D. Little, Marion M. Nau,
Desmond N. Carney, Adi F. Gazdar
& John D. Minna

NCI-Navy Medical Oncology Branch, National Cancer Institute,
National Institutes of Health, and Naval Hospital, Bethesda,
Maryland 20814, USA

Genetic changes involving the *c-myc* oncogene have been observed in human tumours. In particular, the *c-myc* gene is translocated in Burkitt's lymphoma¹⁻³ and is amplified in the human promyelocytic leukaemia cell line, HL-60, which contains double minute chromosomes (DMs)^{4,5}. More recently, an amplified *c-myc* gene has been positioned on a chromosomal homogeneous staining region (HSR) in a human colon cancer cell line, COLO 320, with neuroendocrine properties⁶. Furthermore, *c-myc* is expressed in increased amounts in some human tumour lines, and in some cases, human small cell lung cancers (SCLC) contain DMs and HSRs^{7,8}. These findings prompted us to study the *c-myc* gene and its RNA expression in a series of human lung cancer cell lines. We now report amplification and expression of the *c-myc* oncogene in a system other than B-cell lymphomas⁹, namely human lung cancer. Of 18 human lung cancer cell lines tested, 8 showed an amplified 12.5-kilobase (kb) *EcoRI* *c-myc* DNA band. Of particular interest are five SCLC lines with a high degree of *c-myc* DNA amplification (20-76-fold) and greatly increased levels of *c-myc* RNA. All five lines reside in the variant class of SCLC (SCLC-V) characterized by altered morphology, lack of expression of some SCLC-differentiated functions and more malignant behaviour than pure SCLC^{10,11}. Three of the five lines which have been karyotyped also contain DMs or HSRs. The finding of a greatly amplified *c-myc* gene in all cell lines of the SCLC-V class examined strongly suggests a role for the *c-myc* gene in the phenotypic conversion and malignant behaviour of human lung cancer.

Our laboratory has established and characterized several human lung cancer cell lines for their histological, amine precursor uptake and decarboxylation (APUD) and karyotypic properties (Table 1)¹⁰. There are two major classes of lung cancer: non-small cell lung cancers (NSCLC) which histologically can be further divided into adenocarcinomas, large cell carcinomas and epidermoid carcinomas; and SCLC ('oat cell' carcinomas). The SCLC class can also be divided into pure SCLC and morphological/biochemical variants of small cell lung cancers (SCLC-V)^{10,12-14}. The SCLC-Vs have different cyto-

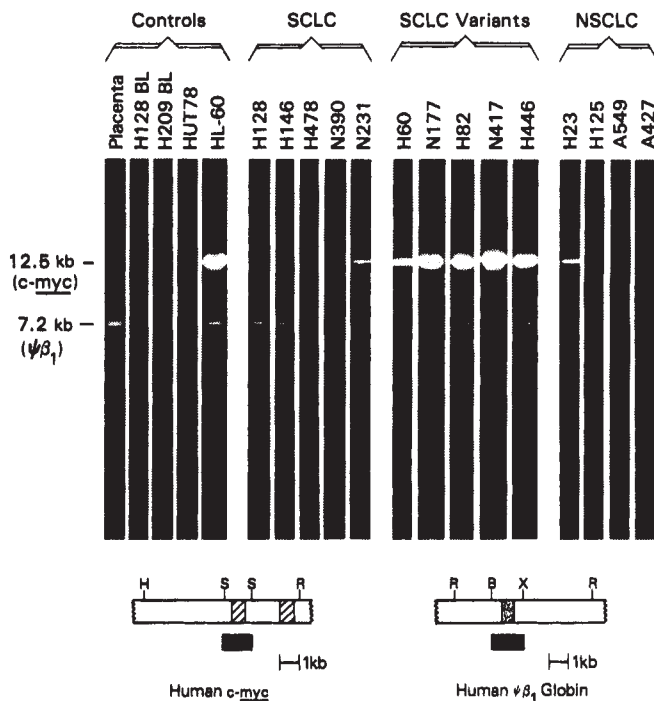


Fig. 1 Hybridization comparison of *EcoRI* digests of representative controls, SCLC, SCLC-V and NSCLC cell line DNAs with human *c-myc* and human pseudo β_1 ($\psi\beta_1$) globin probes.

Methods: Cells were grown as described previously²⁷. DNA was prepared²⁸, 15 μ g digested with *EcoRI*, electrophoresed on a 0.8% agarose gel, denatured and transferred to nitrocellulose essentially as described by Southern²³. Hybridization was performed using 50% dextran sulphate²⁹ with each ³²P-labelled probe indicated on the schematic map and washed at 52 °C as described previously¹. The 1.6-kb *SstI* human *c-myc* probe was obtained from a 12.7-kb *EcoRI* fragment isolated and cloned from human lymphocyte DNA³⁰. Human $\psi\beta_1$ globin is a 1.7-kb *BglII*, *XbaI* fragment obtained from clone λ H γ G4 as previously described³¹. Hatched blocks indicate exons of the *c-myc* probe; stippled blocks indicate the $\psi\beta_1$ globin gene; solid blocks show specific probe used for hybridization. B represents *BglII* site; R, *EcoRI* site; H, *HindIII* site; S, *SstI* site; X, *XbaI* site. Both probes are shown in the 5' to 3' orientation.

logical and histological features from histologically pure SCLC, fail to express typical SCLC APUD markers such as L-dopa decarboxylase (EC 4.1.1.28) (Table 1) or peptide hormones, and have no neurosecretory granules^{10,11,15,16}. However, in contrast to NSCLC, the SCLC-Vs exhibit the SCLC properties of elevated levels of the brain form (BB isoenzyme) of creatine kinase (EC 2.7.3.2)¹¹ and neurone-specific enolase (EC 4.2.1.11)¹⁷. We have also found that the SCLC-Vs have a faster doubling time, higher cloning efficiency, increased tumorigenicity in athymic nude mice, and are more resistant to X-rays than their pure small cell counterparts^{10,11,18}. Clinical correlates of these properties have also been noted. Patients who have the SCLC-V in their diagnostic biopsy specimens have a poorer prognosis than patients with histologically pure SCLC^{12,19}. By these criteria, SCLC-Vs may be considered to possess a more malignant phenotype than pure SCLC.

Previously, we have noted that cell lines established from SCLC-V tumours contain DMs and HSRs (Table 1)⁷, and the presence of these features has been associated with gene amplification^{20,21}. Because of these cytogenetic, clinical and cell biological characteristics, we have previously speculated that the SCLC-V tumours might have amplified genes related to growth, malignant behaviour or therapy resistance²².

We analysed *EcoRI*-digested genomic DNA from 18 well characterized human lung cancer lines with both a human *c-myc* probe and a human pseudo- β_1 ($\psi\beta_1$) globin probe using the

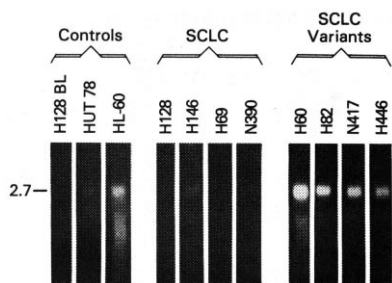


Fig. 2 Hybridization of control and lung cancer cell line RNAs to the 1.6-kb *SstI* human *c-myc* probe.

Methods: A postnuclear supernatant prepared from fresh cell cultures²⁷ essentially as described by Gorman *et al.*³² was phenol/chloroform extracted and ethanol precipitated. 10 µg of total cytoplasmic RNA was denatured and electrophoresed on a 1% agarose-formaldehyde gel according to Lehrach³³, modified by using 0.22 M formaldehyde in the gel and electrophoresing at 160 V for 4 h. The gel was soaked in 3 M NaCl, 0.3 M Na-citrate; transfer, hybridization and washing are essentially as described for DNA.

Southern nitrocellulose transfer method²³ (see Fig. 1 legend for the probe descriptions). The nitrocellulose filters were hybridized with the *c-myc* probe and the $\psi\beta_1$ probe and autoradiography was performed. The single-copy $\psi\beta_1$ globin gene acts as an internal control for the amounts of DNA transferred to the filter and is used to estimate copy number of the *c-myc* gene from one cell line to another.

Representative results are seen in Fig. 1. In all DNAs tested, the human *c-myc* probe detected a 12.5-kb *EcoRI* fragment (Fig. 1) which has been shown to be the germ-line or non-rearranged human *c-myc* gene fragment^{1,2}. Relative to the $\psi\beta_1$ signal, seven NCI cell lines (N231, H60, N177, H82, N417, H446, H23) showed an intense signal corresponding to the 12.5-kb *EcoRI* *c-myc* fragment, indicating that they are amplified in this fragment (Fig. 1). In several *c-myc* amplified cell line DNAs, we also detected *c-myc* hybridizing bands other than the germ-line band. Two lines (H60, N417) each show a unique, additional amplified band (16 and 14 kb) whereas H82 has multiple *c-myc* hybridizing bands. Analyses of additional DNA preparations of these same cell lines also reveal *c-myc* hybridizing bands. These non-germ-line bands are consistent with possible *c-myc* rearrangement and are under further study. We cannot rule out the possibility that these non-germ-line bands represent polymorphisms of the *c-myc* gene because we have been unable to study normal tissue from these patients. However, no *c-myc* polymorphisms have been reported.

Quantitative comparison of the *c-myc* signal with the $\psi\beta_1$ globin signal by densitometric analysis revealed: a 24-fold *c-myc* amplification in HL60 DNA in accord with previous reports^{4,5}; a 20–76-fold amplification in SCLC-V lines (H60, N177, H82, N417, H446); a 4–6-fold amplification in two of the SCLC lines (N390, N231); and a 20-fold increase in NSCLC line H23 (Table 1). These estimates were confirmed using the dot-blot hybridization technique by serially diluting DNAs, spotting on nitrocellulose paper, hybridizing with the *c-myc* probe, then cutting out and counting individual dots (data not shown).

We wished to determine if the amplification of the *c-myc* gene was reflected in increased *c-myc* RNA levels and to compare these values in SCLC and the SCLC-V lines. Total cytoplasmic RNA was prepared from a series of SCLC and SCLC-V cell lines and analysed by Northern blotting. The human *c-myc* probe detected one 2.7-kb RNA band in each cell line (Fig. 2). This size agrees with previous reports for the human 2.7-kb *c-myc* transcript found both in normal and malignant cells²⁴. Densitometer scans of the *c-myc* RNA signal on the Northern blot autoradiograms were performed, allowing quantitative comparison of the amount of the *c-myc* 2.7-kb RNA signal between cell lines. The four amplified *c-myc* DNA cell lines studied (H60, H82, N417, H446) showed values 15–35-fold

Table 1 Properties of human lung cancer cell lines

Source of DNA	DDC level	Presence of DMs or HSRs	Relative amount of <i>c-myc</i> : DNA RNA	
Controls				
Human placenta	ND	ND	1.0	ND
NCI-H128BL	<0.1	—	0.9	2.3
NCI-H209BL	<0.1	—	1.9	ND
HUT-78	<0.1	—	2.3	5.0
HL-60	ND	DMs	24	18
SCLC				
NCI-H128	174	—	0.5	1.1
NCI-H146	134	—	1.0	5.9
NCI-H478	70	ND	1.1	ND
NCI-N390	479	—	3.8	0.9
NCI-N231*	88	ND	5.9	ND
NCI-H209	233	—	0.8	ND
NCI-H69	264	DMs	1	1.0
NCI-H187	22	—	0.7	ND
SCLC-V				
NCI-H60†	<0.1–9.3	DMs	38	35
NCI-N177†	<0.1	ND	76	ND
NCI-H82	0.6	HSR	25	24
NCI-N417*	<0.1	HSR	47	25
NCI-H446	<0.1	ND	20	15
NSCLC				
NCI-H23	<0.1	—	20	ND
NCI-H125	<0.1	—	1.4	ND
A549	<0.1	—	1.4	ND
A427	<0.1	ND	1.0	ND
NCI-H157	<0.1	—	0.8	ND

In cell line nomenclature, the suffix BL represents B-lymphoblastoid lines derived from SCLC patients; HUT-78 is a human T-cell line; HL-60 is a human promyelocytic leukaemia cell line known to be amplified for *c-myc*^{4,5}. In naming NCI tissue culture cell lines the prefix H represents a line grown in tissue culture without passage through a nude mouse; the prefix N represents a cell line passaged from a patient or tissue culture line into a nude mouse and from the heterotransplant into tissue culture. DDC represents L-dopa decarboxylase specific activity (nanomoles of CO₂ released from [¹⁴C]-L-dopa per h per mg protein assayed as described elsewhere^{15,26}. ND, not determined. In column 3, — indicates no DMs or HSRs present. To determine the relative amount of *c-myc* in DNA, densitometer tracings (Beckman DU-8 spectrophotometer) were made of the 12.5-kb *c-myc* band and the 7.2-kb $\psi\beta_1$ band on autoradiograms exposed for varying periods of time depending on band intensity so as to be within the linear response range of the film. The area under the peak for the *c-myc* band was divided by the area under the peak for the $\psi\beta_1$ band and the values were normalized to give a value of 1.0 for placenta DNA. Relative amounts of *c-myc* RNA were determined by making densitometer tracings (DU-8 spectrophotometer) of the 2.7-kb *c-myc* RNA band from autoradiograms of differing exposure times so as to be within the linear response range of the film. The area under the peak for the *c-myc* RNA band was divided by the *c-myc* peak area of SCLC line H69 and the values were normalized to give a value of 1.0 for H69.

* N417 is a SCLC-V cell line which developed from N231, a pure SCLC cell line, after growth in tissue culture for 6 months.

† H60 (late passage) became a SCLC-V after 2 yr in tissue culture. N177 is an independently maintained nude mouse heterotransplant started from the same original tumour which developed into a SCLC-V following 2 yr of heterotransplantation.

over unamplified SCLC lines (H128, H69) (Table 1). One non-amplified SCLC cell line, H146, expressed sixfold more *c-myc* RNA than these non-amplified lines, while one of the minimally amplified SCLC lines (N390) expressed *c-myc* RNA equivalent to the SCLC levels.

Although no consistent abnormality of chromosome region 8q (the location of the human *c-myc* oncogene^{1,25,34}) has been noted in either the SCLC or SCLC-V lines, cell lines H82 and N417 both have an HSR on chromosomes other than 8, while line H60 has DMs²². Further studies are necessary to determine

whether the germ-line *c-myc* gene is amplified within the HSRs in these lines.

As we have not tested SCLC-V specimens taken directly from a patient, we do not know if the *c-myc* amplification occurred in the patient or during growth in cell culture. However, SCLC-V cell lines H82 and H446 were both established from biopsies showing the SCLC-V phenotype and contained amplified *c-myc* genes in the earliest samples tested (passage 6), suggesting that the *c-myc* amplification did occur in the patient.

Two important questions remain unanswered. First, what is the relationship of tumour cells of the SCLC phenotype to those of the SCLC-V class? That is, do SCLC cells 'convert' to SCLC-V or do SCLC-V cells represent an outgrowth of a subpopulation present in the original tumour? Second, does *c-myc* gene amplification cause the phenotype associated with the SCLC-V class? The first question awaits cloning studies. The second question can be directly approached by inserting multiple copies of the *c-myc* gene into cells of the SCLC type.

Although much attention has focused on the role of *c-myc* in human B-cell lymphomas⁹, the *c-myc* oncogene may also be involved in other human malignancies¹⁻⁶. Altogether, our results suggest that changes in the number, organization and expression of the *c-myc* gene are common in human lung cancer lines. This correlation of *c-myc* amplification and expression with the clinical and biological behaviour of the SCLC-V class suggests that *c-myc* genes may be of clinical relevance in these non-lymphoid tumours.

We thank Philip Leder and James Battey for giving us their human *c-myc* clone before publication; Arthur W. Nienhuis for the $\psi\beta_1$ globin probe; J. Michael Bishop for communicating results before publication; Sylvia Stephenson, Ed Russel, Alfreda Simmons and Virginia Bertness for their tissue culture work; and W. Michael Kuehl for critical review of the manuscript. The opinions and assertions presented herein are those of the authors and do not necessarily reflect the official view of the US Department of Defense.

Received 13 July; accepted 20 August 1983.

1. Taub, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7837-7841 (1982).
2. Dalla-Favera, R. *et al. Science* **219**, 963-967 (1983).
3. Erikson, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 820-824 (1983).
4. Dalla-Favera, R., Wong-Staal, F. & Gallo, R. C. *Nature* **299**, 61-63 (1982).
5. Collins, S. & Groudine, M. *Nature* **298**, 679-681 (1982).
6. Altalo, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 1707-1711 (1983).
7. Whang-Peng, J. *et al. Cancer gen. Cytogenet.* **6**, 119-134 (1982).
8. Whang-Peng, J. *et al. Science* **215**, 181-182 (1982).
9. Klein, G. *Cell* **32**, 311-315 (1983).
10. Gazdar, A. F., Carney, D. N., Guccion, J. G. & Balyin, S. B. in *Small Cell Carcinoma of the Lung* (eds Greco, F. A., Oldham, R. K. & Bunn, P. A.) 145-175 (Grune & Stratton, New York, 1981).
11. Gazdar, A. F. *et al. Cancer Res.* **41**, 2773-2777 (1981).
12. Abelloff, M. D. *et al. Am. J. Med.* **66**, 757-762 (1979).
13. Matthews, M. J. in *Lung Cancer: Progress in Therapeutic Research* (eds Muggia, F. & Rozencweig, M.) 155-165 (Raven, New York, 1979).
14. Matthews, M. J. & Gazdar, A. F. in *Lung Cancer I* (ed. Livingston, R. B.) 283-306 (Martinus Nijhoff, Boston, 1981).
15. Baylin, S. B. *et al. Cancer Res.* **40**, 1990-1996 (1980).
16. Moody, T. W. *et al. Science* **214**, 1246-1248 (1981).
17. Marangos, P. J., Gazdar, A. F. & Carney, D. N. *Cancer Lett.* **15**, 67-71 (1982).
18. Carney, D. N., Mitchell, J. B. & Kinsella, T. J. *Cancer Res.* **43**, 2806-2811 (1983).
19. Radice, P. A. *et al. Cancer* **50**, 2894-2902 (1982).
20. *Gene Amplification* (ed. Schimke, R. T.) 1-339 (Cold Spring Harbor Laboratory, New York, 1982).
21. Curt, G. A. *et al. New Engl. J. Med.* **308**, 199-202 (1983).
22. Whang-Peng, J. *et al. in Gene Amplification* (ed. Schimke, R. T.) 107-113 (Cold Spring Harbor Laboratory, New York, 1982).
23. Southern, E. *J. molec. Biol.* **98**, 503-517 (1975).
24. Eva, A. *et al. Nature* **295**, 116-119 (1982).
25. Neel, B. G., Jhanwar, S. C., Chaganti, R. S. K. & Hayward, W. S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7842-7846 (1982).
26. Beaven, M. A., Wilcox, G. & Terpestra, G. K. *Analytic. Biochem.* **84**, 648-641 (1978).
27. Gazdar, A. F. *et al. Cancer Res.* **40**, 3502-3507 (1980).
28. Hieter, P. A. *et al. Nature* **294**, 536-540 (1981).
29. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).
30. Battey, J., Moulding, C., Taub, R. & Leder, P. *Cell* (submitted).
31. Fritsch, E. F., Lawn, R. M. & Maniatis, T. *Cell* **19**, 959-972 (1980).
32. Gorman, C. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 677-6781 (1982).
33. Lehrach, H. *et al. Biochemistry* **16**, 4743-4751 (1977).
34. Dalla-Favera, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7824-7827 (1982).

Are there two classes of VSG gene in *Trypanosoma brucei*?

John R. Young, E. Nancy Miller,
Richard O. Williams* & Mervyn J. Turner

Medical Research Council Biochemical Parasitology Unit, Molteno Institute, Downing Street, Cambridge CB2 3EE, UK

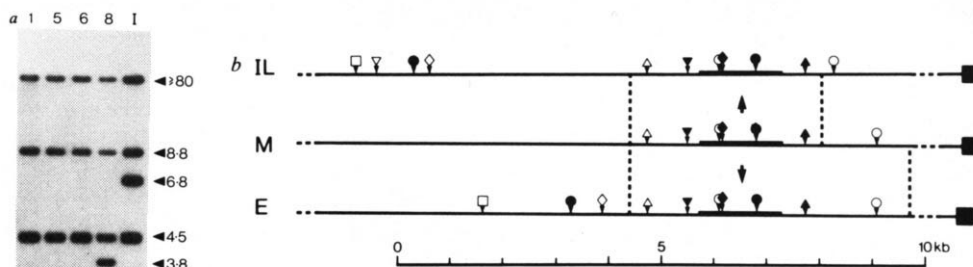
* International Laboratory for Research on Animal Diseases, PO Box 30709, Nairobi, Kenya

Antigenic variation in the African trypanosomes involves the sequential expression of genes coding for different variant surface glycoproteins (VSGs) (reviewed in refs 1-3). When expression of some VSG genes is switched on, a newly duplicated copy of the expressed gene has been observed within the trypanosome genome, which is not found after the gene's expression is switched off again¹⁻¹¹. The duplicated copy has therefore been called an expression-linked copy (ELC). The expression of the gene appears to be strictly coupled to the presence of the ELC. This has led to the hypothesis that the duplicative transposition generating the ELC may itself be responsible for the control of VSG expression^{1,9,10,12,13}. With other VSG genes, expression-linked duplication has not been observed, and expression is clearly not controlled in this way^{1-3,14-19}. Data are presented here which demonstrate that either of these observations may be obtained with a single VSG gene, depending on the chance selection of particular clones from antigenically switched populations. Thus, the different observations do not imply the existence of two distinct classes of VSG gene controlled by different mechanisms, but different aspects of processes common to all VSG genes.

Antigenic switching in trypanosomes occurs at a low frequency (about 10^{-4} events per cell generation)^{20,21}. Infection of an immunocompetent host with an antigenically homogeneous population produces a relapsing parasitaemia in which the first peak contains trypanosomes of the inoculated antigenic type. The first relapse population is antigenically heterogeneous, generated from a small number of organisms that switched early in the infection^{3,20,21}. Previous studies of the genomic rearrangements associated with antigenic variation have compared the genome of the cloned population used to initiate an infection with that of a small number of clones isolated from the relapse population. Since both the emergence of the relapse population and cloning from it are highly selective events, there is no guarantee that the genomes compared represent those of the individual organism immediately before and after the switch event. One object of the study reported here was to determine whether these limitations of the experimental system might account for the differences in observations that have been obtained with different VSG genes.

Analysis of the ILTat 1.3 VSG gene in a series of closely related clones within the ILTAR 1 serodeme showed that the modulation of expression of this gene did not involve duplication¹⁶. There were five copies of the gene in all the clones studied, two of which are not distinguishable in most Southern blot experiments. Two copies were telomeric, one of which was shown, by demonstration of expression-specific DNase I sensitivity, to be the copy expressed in the ILTat 1.3 trypanosome clone. It occupied a site very similar to that occupied by the ELCs of other VSG genes. The other telomeric copy was located at the end of a minichromosome²². Analysis of the same gene in the genetically similar ETAR 1 trypanosome series revealed that most of the ETAR 1 clones contained four of the five gene copies present in the ILTAR 1 clones, but lacked the expressed copy. A new telomeric copy appeared, however, in the clone ETAT 1.8 that expresses a VSG serologically identical to ILTat 1.3, which had arisen by duplication of the minichromosome copy²³. This situation is summarized in Fig. 1. Although this

Fig. 1 Newly duplicated copy of the ILTat 1.3 gene in ETat 1.8. **a**, The Southern blot shows the hybridization of a 450-base pair (bp) 5' ILTat 1.3 cDNA probe to *Hind*III-digested DNA from four ETAR 1 trypanosome clones and one ILTAR 1 clone. The derivation of these clones is described elsewhere^{23,24}. The ETat clones express VSG ETat 1.x where x is the number shown above the track. The ILTat clone (I) expresses ILTat 1.4. ETat 1.8 is serologically identical to ILTat 1.3. DNA isolation, preparation of the probe and procedures for the Southern blotting were exactly as described elsewhere¹⁴. *Hind*III cuts once, 1,140 bp from the 5' end of the ILTat 1.3 cDNA sequence²⁵, so that each copy of the gene produces a single 5' flanking sequence that will hybridize to this probe¹⁶. The 5' fragments from two of the non-telomeric copies are identical (4.5 kb, ref. 16). The >80-kb fragment is from the minichromosome copy, and the 8.8-kb fragment from the other non-telomeric copy. The 3.8-kb band in ETat 1.8 is from the new copy. **b**, The restriction maps are of the new copy in ETat 1.8 (E), the expressed copy in ILTat 1.3 (IL) and the common 'minichromosome' copy (M) from which both appear to have been derived by duplication. Solid blocks are at the 3' end, and represent the chromosome ends. The thickened region is the position of the cDNA sequences. Enzymes are: ◆, *Ava*I; ▲, *Ava*II; ○, *Bst*EII; △, *Hinc*II; ●, *Hinc*II; ◇, *Msp*I; □, *Eco*RI; ▽, *Eco*RV; ▼, *Hinf*I. Dashed lines are possible delimiters of the duplicated regions^{16,23}. Construction of the maps has been described elsewhere²³.



gene was controlled without duplication in the ILTAR 1 clones, it appeared that expression-linked duplication occurred in the ETAR 1 series.

To avoid the selection implicit in cloning from relapse populations, DNA was isolated directly from trypanosomes in the first relapse parasitaemia following infection of rats with approximately 10 organisms from a homogeneous (>99.5%) population of the ETat 1.8 clone. The relapse populations contained less than 1 individual per 1,000 expressing ETat 1.8 VSG. Figure 2 shows an analysis of the DNA from relapse populations in two rats, with a cDNA probe that detects a single 5' flanking fragment from each of the distinguishable gene copies in these *Hind*III digests (tracks 1–5). The 3.8-kilobase (kb) fragment from the newly duplicated gene copy in ETat 1.8 is also detected in both relapse populations. In the first (R1) the relative intensity of this band indicates that most of the trypanosomes in this relapse retained the newly duplicated copy. In the second (R2) the relative intensity is about one-third of the same band in the ETat 1.8 DNA. Further propagation of the relapse population (R2') through about 25 cell generations did not cause significant diminution of the band's relative intensity.

The emergence of a relapse population following the host's immune response is a highly selective process³. The population is probably derived from a small number of organisms which switch antigen expression early in the infection and therefore outgrow others switching later. The genomic composition of the relapse population can be expected to reflect any heterogeneity among this small number of organisms, together with any subsequent changes during their expansion into the relapse population. In the experiments reported here this expansion involved approximately 30 cell generations. Because a further 25 cell generations of the R2 population did not further reduce the intensity of the 3.8-kb band, its reduction is more likely to be due to heterogeneity in the initially switched organisms than to gradually accumulating changes. Thus, it appears that in this rat, two-thirds of the switching events producing the relapse population were associated with loss of the newly duplicated copy. Whether this occurred during or closely after the switching event is not clear. In the other third, and in most of the events in the other rat, switching did not involve loss of the new gene copy.

Of three clones which were derived from the heterogeneous relapse population (R2), two lacked the 3.8-kb fragment entirely, while the third contained this fragment with the same stoichiometry as in the ETat 1.8 clone (Fig. 2, tracks 6–8). None of the relapse clones expressed the ETat 1.8 VSG. None of the sites shown in Fig. 1, with the exception of the distance to the chromosome end, were different in the retained gene copy in clone C3. The 3.8 kb *Hind*III fragment was selectively sensitive

to DNase 1 digestion in the ETat 1.8 clone, but not in clone C3. Thus, it is unlikely that the retained sequences are expressed as part of a VSG gene altered by internal recombination. Expansion of the clone which maintained the 3.8-kb fragment involved approximately 90 cell generations before the DNA was isolated, but the fragment remained in most of the progeny. This again suggests that presence or absence of the fragment is the result of alternative courses at or close to the switching event, and confirms that the reduced intensity of the band in the R2 population was due to heterogeneity of gene content among different trypanosomes.

All previously studied clones from the ILTAR 1 serodeme contain the ILTat 1.3 gene copy which is expressed in the original ILTat 1.3 clone, whether they express ILTat 1.3 or not^{14,16}. This copy produces the 6.8-kb band in the *Hind*III digests of ILTat DNA shown in the right half of Fig. 2. The clone MIAG227 (13'), which also expresses ILTat 1.3, was derived, after tetse fly passage, from the same member of the ILTat series that preceded the isolation of the original ILTat 1.3 clone²⁰. It contains the same 6.8-kb 5' flanking *Hind*III fragment (Fig. 2), as well as 5' flanking fragments with other enzymes (not shown), that are characteristic of the expressed copy in that particular site. Thus, it appears that this gene remained present during the derivation of this clone. In Fig. 2, tracks 12–15 contain DNA from relapse populations (R1–R6) derived from MIAG227²⁰. One of these (R4) contains the same complement of the 6.8-kb fragment as the MIAG227 DNA; the others contain less. A clone derived from the R6 population did not contain this gene copy (Fig. 2, track 16). These results were confirmed with other restriction enzymes. Thus, as with the ETat clones, loss of this gene copy may, but does not necessarily, occur close to or during the antigenic switching event.

If we had examined only ETat 1.8, its immediate precursor ETat 1.6 and the first two clones from its relapse, the typical ELC phenomenon would have been observed. However, the third clone and the persistence of the 'ELC' in a substantial proportion (R2) or the whole (R1) of the relapse populations demonstrate that such data cannot justify the conclusion that the presence of the duplicated gene copy and its expression are strictly linked. This must apply equally to comparable observations made with other VSG genes⁹.

VSG genes have been classified into two groups according to whether or not expression-linked duplication was apparently observed¹. This division allowed that the duplicative transposition involved in ELC generation, with simultaneous displacement of the pre-existing VSG gene, at a single expression site, might constitute the mechanism of mutually exclusive VSG expression, despite the evidence that other VSG genes were not controlled in this way^{1,9,10,12}. The data presented here show

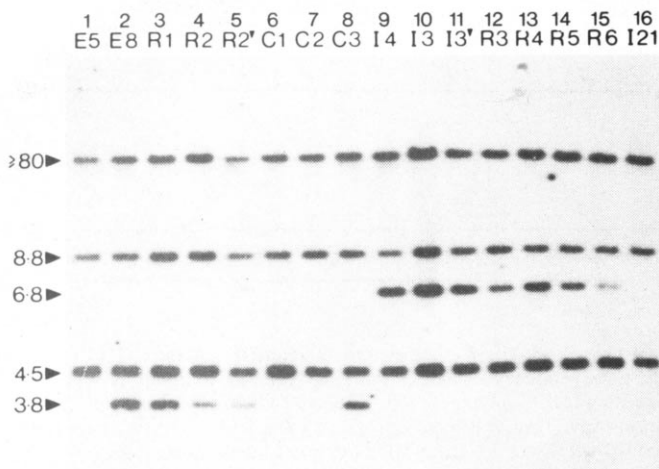


Fig. 2 Southern blot analysis of relapse populations of ETat 1.8 and MIAG227 for content of ILTat 1.3 VSG genes. DNA was isolated from: ETat 1.5 (E5), ETat 1.8 (E8), relapse populations from ETat 1.8 (R1 and R2), three clones derived from R2 (C1–C3), ILTat 1.4 (I4), ILTat 1.3 (I3), MIAG227 (I3'), also expressing ILTat 1.3 VSG), relapse populations from MIAG227 (R3–R6) and a clone (I21'), expressing ILTat 1.21) derived from R6. Sizes of fragments are indicated in kilobase pairs.

Methods: The MIAG227 relapses have been described elsewhere²⁰. DNA was isolated from these after one 3-day passage in mice of the cryopreserved populations. Infections in two rats (R1, R2) were initiated by intraperitoneal injection of approximately 10 trypanosomes from a homogeneous population of ETat 1.8. The peak of the first parasitaemia was on days 6–7. No parasites were detected on day 8. On day 11 the relapse parasitaemia had reached approximately 10⁹ per ml and DNA was isolated from these trypanosomes. The cryopreserved R2 population was passaged for 3 days in a mouse (10⁶ inoculated) and 3 days in a rat (10⁷ inoculated) before again isolating DNA (R2+). Clones C1–C3 were obtained after expansion of the cryopreserved day 11 population (R2) in mice. The ILTat 1.3 and ILTat 1.4 clones have been described elsewhere (clones C1 and D1 in ref. 14). Procedures for immunofluorescence²⁰ and for DNA isolation and Southern blotting were as described previously¹⁴. Intensities of bands in this autoradiogram were estimated by densitometric scanning and integration of peak areas.

that expression of a single VSG gene may involve either the activation of a newly duplicated copy (ETat) or the activation of a pre-existing telomeric copy (ILTat)¹⁶. In either case, whether the expressed copy is seen to be lost when expression is switched off may depend solely on the chance selection of particular clones from the relapse population. Thus, results characteristic of either of the VSG gene classes may be obtained with a single VSG gene. This implies that the classification of VSG genes on the basis of different experimental results does not necessarily reflect inherent differences in the processes affecting particular genes. The two kinds of experimental observation should be interpreted as reflecting different aspects of processes common to all VSG genes rather than specific properties of particular genes.

The generation of ELCs involves duplication of the basic copy VSG gene and transposition of the duplicated segment into a telomeric receptor site^{6,7,9,10,12}. The finding of an identical array of restriction enzyme sites beyond the 5' ends of ELCs of different VSG genes suggested that the receptor site might be the single expression site required for mutually exclusive expression^{1,6,10}. It is now evident that the expressed copies of VSG genes for which expression-linked duplication had not previously been observed probably arise through an analogous duplicative transposition^{15,16}. A number of such duplicated genes may, however, co-exist in different sites in the genome while only one of them is expressed. When these data are considered together, as required by the above conclusions, duplicative transposition described for ELCs may be a pre-

quisite for the eventual expression of all VSG genes, but a different mechanism must be responsible for ensuring mutually exclusive expression.

We thank Dr G. Matthysens and Dr J. J. Doyle for critical comments on the manuscript. J.R.Y. was supported by the International Laboratory for Research on Animal Diseases (ILRAD), Nairobi. M.J.T. received support from the WHO/World Bank/UNDP Special Program for Research and Training in Tropical Diseases. E.N.M. is a Beit Memorial Research Fellow.

Received 13 July; accepted 7 September 1983.

1. Borst, P. & Cross, G. A. M. *Cell* **29**, 291–303 (1982).
2. Englund, P. T., Hadjuk, S. L. & Marini, J. C. A. *Rev. Biochem.* **51**, 695–726 (1982).
3. Turner, M. J. *Adv. Parasit.* **21**, 69–153 (1982).
4. Hocjmakers, J. H. J., Frasch, A. C. C., Bernards, A., Borst, P. & Cross, G. A. M. *Nature* **284**, 78–80 (1980).
5. Agabian, N., Thomashow, L., Milhausen, M. & Stuart, K. *Am. J. trop. Med. Hyg. Suppl.* **29**, 1043–1049 (1980).
6. Pays, E., Lheureux, M. & Steinert, M. *Nucleic Acids Res.* **10**, 3149–3163 (1982).
7. Pays, E., Lheureux, M. & Steinert, M. *Nucleic Acids Res.* **10**, 4225–4238 (1981).
8. Michels, P., Bernards, A., Van der Ploeg, L. & Borst, P. *Nucleic Acids Res.* **10**, 2353–2366 (1982).
9. Van der Ploeg, L. H. T., Bernards, A., Rijsewijk, F. A. M. & Borst, P. *Nucleic Acids Res.* **10**, 593–609 (1982).
10. Longacre, S. *et al. Molec. cell. Biol.* **3**, 399–409 (1983).
11. Bernards, A. *et al. Cell* **27**, 497–505 (1981).
12. Laurent, M. *et al. Nature* **302**, 263–266 (1983).
13. Young, J. R., Donelson, J. E., Majiwa, P. A. O., Shapiro, S. Z. & Williams, R. O. *Nucleic Acids Res.* **10**, 3149–3163 (1982).
14. Donelson, J. E., Young, J. R., Majiwa, P. A. O. & Williams, R. O. *Nucleic Acids Res.* **10**, 6581–6595 (1982).
15. Young, J. R., Shah, J. S., Matthysens, G. & Williams, R. O. *Cell* **32**, 1149–1159 (1983).
16. Williams, R. O., Young, J. R., Majiwa, P. A. O., Doyle, J. J. & Shapiro, S. Z. *Cold Spring Harb. Symp. quant. Biol.* **45**, 945–949 (1980).
17. Borst, P. *et al. Cold Spring Harb. Symp. quant. Biol.* **45**, 1033–1036 (1980).
18. Borst, P. *et al. J. Cell. Biochem.* (in the press).
19. Miller, E. N. & Turner, M. J. *Parasitology* **82**, 63–80 (1981).
20. Van Meirvenne, N., Janssens, P. G. & Magnus, E. *Ann. Soc. Belg. Med. trop.* **55**, 1–23 (1975).
21. Williams, R. O., Young, J. R. & Majiwa, P. A. O. *Nature* **299**, 417–421 (1982).
22. Young, J. R., Turner, M. J. & Williams, R. O. *J. cell. Biochem.* (in the press).
23. Lumsden, W. H. R. & Herbert, W. J. *Trans. R. Soc. trop. Med. Hyg.* **69**, 205–208 (1975).
24. Rice-Ficht, A. C., Chen, K. K. & Donelson, J. E. *Nature* **298**, 676–679 (1981).

Extrachromosomal replication of *copia*-based vectors in cultured *Drosophila* cells

J. H. Sinclair & J. H. Sang

Department of Biological Sciences, University of Sussex, Falmer, Brighton, Sussex BN1 9QG, UK

J. F. Burke* & D. Ish-Horowicz

Laboratory of Developmental Genetics, Imperial Cancer Research Fund, Burtonhole Lane, London NW7 1AD, UK

The functional analysis of cloned *Drosophila* genes *in vivo* has so far not been possible because of the lack of an appropriate cultured cell transformation system. We have developed a transformation method which uses the *Escherichia coli* gene coding for xanthine guanine phosphoribosyl transferase (*gpt*)¹, inserted into a cloned *copia*² transposable element, and a medium selective for transformed cells expressing the *gpt* gene. We show that the resulting plasmids (pCVgpt) propagate extrachromosomally in transformed cells, apparently without integrating into the genome. This suggests that the extrachromosomal *copia* circles found in *Drosophila* tissue-culture cells may also replicate semi-conservatively.

Copia is present at 20–40 copies per haploid genome in flies, and at about 150 copies per cell in *Drosophila* tissue culture cells^{3,4}. It is also found as extrachromosomal circular molecules

* Present address: MRC Cell Mutation Unit, University of Sussex, Falmer, Brighton, Sussex BN1 9QG, UK.

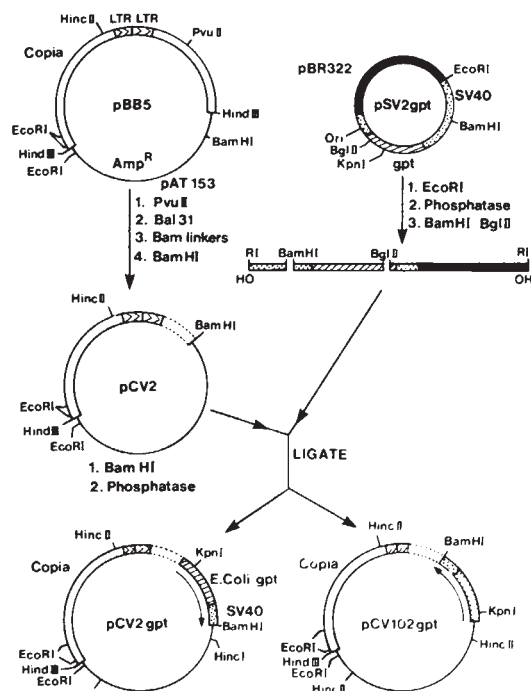


Fig. 1 Construction of the pCVgpt vectors. 3 µg of *PvuII*-digested pBB5 DNA was digested with 1 unit of *Bal31* exonuclease for 3 min at 30 °C. The DNA was ligated to *Bam*HI linkers, cut with *Bam*HI, circularized with *T₄* DNA ligase, and used to transfect DH1 cells using a procedure devised by M. Scott (personal communication). Plasmid pCV2 was derived by omitting the *Bal31* digestion step; the other pCV plasmids were identified from plasmid DNA 'minipreps' and the approximate positions of the *Bam*HI linker determined by *Bam*HI–*Hinc*II digestion. Vectors containing the *gpt* gene were made from suitable pCV plasmids by inserting the *Bam*HI–*Bgl*II fragment of pSV2gpt¹. Dephosphorylation was with calf intestinal phosphatase. pCVgpt vector plasmids were identified from *Hind*III + *Bam*HI digests of DNA minipreps, and the orientation of the *gpt* gene was confirmed with *Kpn*I–*Bam*HI digests. Mini and maxi plasmid preparations were as described elsewhere¹². □, *copia* DNA sequences in pBB5²; □, a long terminal repeat (LTR) of *copia*, the arrows indicating the direction of transcription from the *copia* promoters; ▨, the *E. coli gpt* gene; ▨, SV40 sequences containing the small T antigen splice sequences and the poly(A)-addition site. The thin line represents pAT153 sequences¹³. The arrows indicate the direction of *gpt* transcription. pCV4gpt and pCV104gpt are vectors in which the *Bal31* has removed the 5'-LTR completely and left 203 bp of the 3' LTR, against which lies the *gpt* gene. The *gpt* gene in pCV2gpt is aligned with the *copia* promoter; in pCV102gpt, they are oppositely oriented.

the structure of which resembles that of an unintegrated circular retroviral provirus². The integrated *copia* element has a characteristic structure of a central 4.7 kilobase (kb) region flanked by 276 base pair (bp) direct terminal repeats (LTRs). Two major transcripts, which make up about 2% of the total polyadenylated RNA, are initiated within the 5' LTR of the element^{5,6}. We therefore constructed vectors with the *E. coli gpt* gene inserted downstream from the LTRs (Fig. 1) in the expectation that this would also be transcribed.

The vectors are based on plasmid pBB5, a cloned circular *copia* element that contains two tandem LTRs. The *gpt* gene was inserted into the *Pvu*II site of pBB5 (Fig. 1), 545 bp upstream of the nearer LTR, either aligned with the *copia* LTRs (pCV2gpt), or in the opposite orientation (pCV102gpt). We also made vectors in which the distance between the LTR and the *gpt* gene was varied using *Bal31* exonuclease (Fig. 1).

Drosophila D1 cells were transfected with both types of vector using the calcium phosphate procedure⁷ (Table 1). Like mammalian cells, wild-type *Drosophila* cells are unable to grow on xanthine as the sole purine source. Cells that had acquired an active *gpt* gene were selected by exposure to a xanthine-based

Table 1 Efficiency of transfected DNAs in transforming D1 cells to *gpt*-expression

Transfection DNA	Transformation efficiency
pCV2gpt	1×10^{-4}
pCV102gpt	1×10^{-4}
pSV2gpt	$< 1 \times 10^{-8}$
Carrier D1 DNA	$< 1 \times 10^{-8}$
pCV4gpt	1×10^{-4}
pCV104gpt	1×10^{-4}

Drosophila D1 cells were routinely maintained in M3 medium⁹. In these conditions, they have a plating efficiency of 50–100%¹⁰. For selective medium M3X, M3 was modified as follows; no yeastolate was added, and the medium was supplemented with 200 mg l⁻¹ lactalbumin hydrolysate together with CuSO₄ 5H₂O (10 µg l⁻¹), FeSO₄ 7H₂O (0.5 mg l⁻¹), ZnSO₄ 7H₂O (0.6 mg l⁻¹), MgSO₄ 7H₂O (0.3 mg l⁻¹), vitamin B₁₂ (0.1 mg l⁻¹), plus Grace's vitamin mix¹¹, and 10% fetal calf serum which had been extensively extracted with Norit-GSX charcoal. This semi-defined medium was made selective for expression of the *gpt* gene by the addition of thymidine (24 mg l⁻¹), adenine (13.4 mg l⁻¹), mycophenolic acid (50 mg l⁻¹), xanthine (150 mg l⁻¹) and 10⁻⁶ M methotrexate (a modification of the medium described previously¹). For transformation, 1×10^6 cells were plated into Nunclon (40 ml) tissue culture flasks. After 24 h to allow the cells to settle and resume growth, 0.5 ml of calcium phosphate precipitate⁷ containing 3–5 µg of plasmid DNA and 15–20 µg of D1 carrier DNA was added directly to the medium in the flasks. After an overnight exposure to the DNA at 25 °C, the medium was removed and the cells washed with M3 medium. An expression time of 3 days was allowed for the cells to recover and express the transfected DNA, after which the cells were washed and placed in the selection medium M3X. If the original flasks were greater than half confluent, cells were replated at $2-3 \times 10^6$ cells per flask, 24 h before selection. The selective medium was replaced every 5–7 days. After 4–6 weeks colonies were visible which grew under selection with a doubling time of 2–3 days. Transformation frequency was scored after 6 weeks in M3X selective media. pCV4gpt and pCV104gpt are described in the legend to Fig. 1.

medium (M3X) in which endogenous purine synthesis is blocked with the antibiotics methotrexate and mycophenolic acid. Transformation of D1 cells with both pCV2gpt and pCV102gpt gives colonies that grow in the M3X medium at a frequency of about 10^{-4} (Table 1). Transformed colonies grow slowly in M3X, with a doubling time of about 3 days. There are no transformants with carrier DNA alone or with pSV2gpt DNA (Table 1).

Four independent cell lines were cloned from cultures transfected with pCV2gpt, and Southern blot analysis of their undigested DNA shows, in each case, that the plasmid is present as extrachromosomal DNA, and not integrated into the genomic DNA. Restriction enzyme digests of this DNA shows that the extrachromosomal DNA is identical to the original pCV2gpt vector (Fig. 2).

Surprisingly, the orientation of the *gpt* gene with respect to the LTRs has no measurable effect on transformation frequency (Table 1). A cloned line (line 1) derived from cells transformed with pCV102gpt was analysed by Southern blot hybridization. The lack of hybridization to high molecular weight chromosomal DNA in Fig. 3a shows that the vector pCV102gpt molecules are again extrachromosomal. Lane 2 shows that undigested DNA from line 1 of the pCV102gpt-transformed cells contains a major species of *gpt* gene containing molecules of size ~5.5 kb and a minor 5 kb species. Lane 1 shows that the extrachromosomal molecules are lost when the cells are grown in the absence of selection before collection.

The *gpt*-containing DNA in the pCV102gpt-transformed cells cannot just be residual input DNA. The extrachromosomal DNA shown in Fig. 3a is ~3 kb smaller than the 8.4 kb vector DNA, having rearranged in the selected cells. Also, restriction enzyme digests of line 1 DNA give different fragments from those expected from the input pCV102gpt DNA, for example the rearranged plasmids have more than one *Bam*HI site (Fig. 3b lane 1).

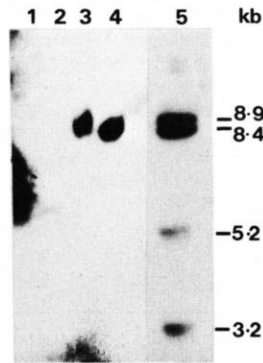


Fig. 2 *Bam*HI and *Eco*RI digests of DNA from a pCV2gpt transformant maintained in selection. DNA from untransformed cells (lanes 1 and 2) or D1 cells transformed with pCV2gpt DNA (lanes 3 and 4) was digested with either *Bam*HI (lanes 1 and 3) or *Eco*RI (lanes 2 and 4), transferred to nitrocellulose and probed¹⁴ with nick-translated pSV2gpt. Markers (lane 5) contain a mixture of *Bam*HI, *Eco*RI and *Bam*HI-*Eco*RI digests of pCV2gpt DNA. *Eco*RI-digested pCV2gpt is about 500 bp smaller than *Bam*HI-digested pCV2gpt (Fig. 1).

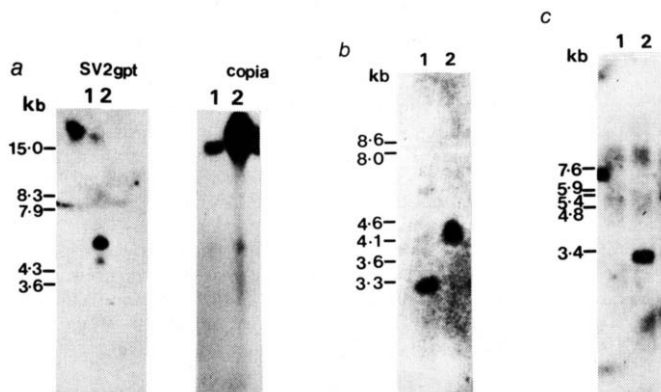


Fig. 3 **a**, Extrachromosomal *gpt* DNA in pCV102gpt transformed cells. Undigested DNA from line 1 cells that had been maintained under selection (lane 2) or in the absence of (lane 1) selection for 6 weeks was analysed by Southern transfer from a 0.5% agarose gel. Filters were first probed with nick-translated pSV2gpt, which was then melted off to allow rehybridization to the *copia* probe pBB5. The size markers were supercoiled plasmid DNAs. **b**, *Bam*HI and *Eco*RI digests of DNA from pCV102gpt transformed cells. DNA from clone 1 cells that were maintained under selection was digested with either *Bam*HI (lane 1) or *Eco*RI (lane 2). The DNA was transferred to nitrocellulose and probed with nick-translated pSV2gpt. **c**, *Bam*HI-*Eco*RI double digests of DNA from pCV102gpt-transformed cells was digested with *Bam*HI and *Eco*RI. DNA from clone 1 cells maintained under selection (lane 2) or in the absence of selection (lane 1) was digested and probed with pSV2gpt.

From the intensity of hybridization, we estimate that there are a maximum of two copies of the *gpt* gene per cell when maintained in the selective medium, and this low copy number may explain the slow doubling of cell numbers by the segregation of cells lacking *gpt*-containing molecules. As might be expected, growth in non-selective medium results in loss of the *gpt* gene. Figure 3a and c show that the *gpt*-containing molecules in selected line 1 (lanes 2) are lost in unselected line 1 (lanes 1). As expected, loss of the *gpt*-containing molecules is accompanied by the loss of the ability to grow on xanthine. Thus, *gpt*-transformed cells grown for several weeks in non-selective medium are no longer able to grow in the selective M3X medium.

The rearrangement of the *gpt*-containing DNA appears only to have occurred in the cells transformed with pCV102gpt, and not in those transformed with pCV2gpt. Although we have not yet shown that the LTR provides a promoter for *gpt* transcrip-

tion, the simplest explanation for the rearrangement is that it is associated with expression of the *gpt* gene. The structure of pCV102gpt provides no obvious promoter for the *gpt* gene, yet its expression is essential for growth in M3X of the pCV102gpt-transformed cells. Rearrangement does not seem to be necessary for replication, as both pCV2gpt and pCV102gpt are able to replicate extrachromosomally in a *shibire*^{ts} cell line (data not shown). In any case, propagation of the rearranged circular *gpt* gene-containing molecules in pCV102gpt-transformed cells that lack integrated *gpt* DNA (Fig. 3) indicates that the vector is replicating extrachromosomally.

The structural similarities between circular *copia* elements and retroviral proviruses⁸ raise the possibility that the *gpt*-containing molecules in transformed cells replicate by reverse transcription of a *copia*-*gpt* RNA transcript into a double-stranded DNA molecule which can circularize. This seems very unlikely because the cells lack integrated copies of the *gpt* gene that could be templates for full-length transcription. Moreover, we have made vectors (for example pCV4gpt and pCV104gpt in Table 1 and Fig. 1) in which the *copia* sequences homologous to the retroviral tRNA-primer binding site are deleted but which retain the ability to transform cells.

As our vectors are based on a cloned extrachromosomal circular *copia* element, we think it likely that the uncloned cellular circles also replicate in growing cells. We are currently attempting to map the replication origin to define the *copia* sequences involved.

Finally, vectors such as pCV2gpt should allow the introduction and analysis of cloned genes in cultured *Drosophila* cells. We have used this system to introduce *Drosophila* heat shock genes into tissue culture cells by co-transformation with a pCVgpt vector (unpublished data).

We thank N. Krauzeiwicz for technical assistance. This work was supported by a grant from the SERC.

Received 5 July; accepted 19 September 1983.

1. Mulligan, R. C. & Berg, P. *Science* **78**, 2072-2076 (1981).
2. Flavell, A. J. & Ish-Horowitz, N. *Nature* **292**, 591-595 (1981).
3. Finnegan, D. J., Rubin, G. M., Young, M. W. & Hogness, D. S. *Cold Spring Harb. Symp. quant. Biol.* **42**, 1053-1063 (1978).
4. Potter, S. S., Borein, W. J., Dunsmuir, P. & Rubin, G. M. *Cell* **17**, 415-427 (1979).
5. Flavell, A. J., Lewis, R., Simon, M. A. & Rubin, G. M. *Nucleic Acids Res.* **9**, 6279-6291 (1981).
6. Schwartz, H. E., Lockett, J. J. & Young, M. W. *J. molec. Biol.* **157**, 49-68 (1982).
7. Huttner, K. M., Barbosa, J. A., Scangos, G. E., Pratchera, D. D. & Ruddle, F. H. *J. Cell Biol.* **91**, 153-156 (1981).
8. Temin, H. *Cell* **21**, 599-600 (1980).
9. Shields, G. & Sang, J. H. *Drosoph. Inform. Ser.* **52**, 161 (1977).
10. Sang, J. H. *Adv. Cell culture* **1**, 125-182 (1981).
11. Grace, T. C. C. *Nature* **195**, 788-789 (1962).
12. Ish-Horowitz, D. & Burke, J. F. *Nucleic Acids Res.* **9**, 2989-2998 (1981).
13. Twigg, A. J. & Sherratt, D. *Nature* **283**, 216-218 (1980).
14. Jeffreys, A. J. *et al.* *Cell* **21**, 555-564 (1980).

Transcription in oocytes of highly methylated rDNA from *Xenopus laevis* sperm

Donald Macleod & Adrian Bird

MRC Mammalian Genome Unit, King's Buildings, West Mains Road, Edinburgh EH9 3JT, UK

The genes for ribosomal RNA exist as multiple copies in the genome. Each repeated unit comprises a region that codes for the 40S rRNA precursor, and a spacer region of uncertain function^{1,2} (Fig. 1a). In *Xenopus laevis* there are about 1,000 copies of the dinucleotide sequence C-G in each repeat unit, and of these about 250 can be tested for the presence of 5-methylcytosine using restriction endonucleases^{3,4}. Most of the detectable C-Gs are heavily methylated, but in somatic cells unmethylated C-Gs occur⁵ in a 60 base pair (bp) sequence (NTS-60) that is repeated in the spacer^{6,7}. In contrast, the spacer of sperm rDNA is heavily methylated at these and all

other testable C-Gs. Loss of methylation at NTS-60 takes place during the first day of embryonic development, near the time when rDNA transcription begins⁴. In an attempt to assess the significance of this developmental change in methylation, we have isolated sperm rDNA and investigated whether it can be transcribed in oocytes. We have found that sperm rDNA is transcribed as efficiently as cloned rDNA, although no loss of methylation was detectable. Direct sequencing of sperm rDNA showed that all 19 C-Gs in the promoter are highly methylated. Thus, in the case of rDNA injected into oocytes, loss of methylation is unnecessary for effective transcription.

Xenopus rDNA is unusual in that it can be purified from genomic DNA by physical methods⁵. A consequent advantage is that the genes are purified with their endogenous pattern of methylation intact. We isolated rDNA from total blood or purified *X. laevis* sperm heads using CsCl/actinomycin D gradients as a first purification step. Levels of C-G methylation were checked by digesting the *Eco*RI fragments of purified rDNA with *Hpa*II and *Ava*I (Fig. 1b). As expected, *Eco*RI fragments of sperm rDNA (lane 3) were resistant to *Hpa*II and *Ava*I as a result of C-G methylation at these sites (lanes 4 and 5). On the other hand, the spacer *Eco*RI fragment of blood rDNA was cut at variable sites corresponding to clusters of NTS-60 in the spacer (lanes 1 and 2). In order to extend our analysis to all C-Gs in the repeat unit of sperm rDNA, we nick-translated sperm rDNA in the presence of [α -³²P]dGTP alone and subjected it to nearest neighbour analysis. This method⁹ measures the extent of C-G methylation from the level of 5-methyl dCMP on the autoradiograph. The result suggested that over 97% of C-Gs in this preparation of sperm rDNA are methylated (see legend to Fig. 2 for limitations of the technique).

About 2 μ g each of *X. laevis* sperm and blood rDNA were digested with *Eco*RI and ligated at low concentration to generate circles. These were then injected into oocytes of a closely related species, *X. borealis*^{10,11}. Maintenance of methylation in the oocyte was tested by recovering nucleic acids from injected oocytes. The injected DNA was detected by blotting¹² and hybridization to a spacer-specific probe of *X. laevis* rDNA, which hybridizes minimally with *X. borealis* rDNA (Fig. 1c). As sperm rDNA remained resistant to *Hpa*II digestion after injection (lanes 3 and 4), we infer that injected DNA was not demethylated during incubation in the oocyte nucleus. Previous studies have drawn similar conclusions^{13,14}.

Transcripts from the injected genes were assayed by S₁-mapping¹⁵ with a single-stranded probe that had been end-labelled with [γ -³²P]ATP and polynucleotide kinase (Fig. 3). This probe hybridizes to a 93 nucleotide fragment of *X. laevis* rRNA, but does not hybridize to *X. borealis* rRNA (Fig. 3, lanes 1 and 2). The results showed that equivalent concentrations of circular sperm or blood rDNA gave equivalent levels of transcription (Fig. 3, lanes 5 and 6).

Moss¹⁶ has shown that competition between different rDNA molecules can reveal differences in transcriptional efficiency that are not detected when clones are tested singly. We therefore co-injected sperm rDNA with a cloned (and therefore unmethylated) rDNA fragment. The clone was a derivative of pXlr108c, the transcription of which has been extensively studied in the oocyte^{10,16}. The transcript from the clone was distinguishable from that of sperm rDNA because of a short deletion within the transcribed region, which gave rise to an S₁-resistant fragment of 61 nucleotides, compared with 93 nucleotides for undeleted rDNA (Fig. 3, lane 7). When sperm or blood rDNA circles were injected at a 1:1 molar ratio with the deleted clone, the levels of transcription from cloned and chromosomal rDNAs were approximately equal (Fig. 3, lanes 3, 4 and 9). When the molar ratio of sperm rDNA to cloned rDNA was increased to 4:1, the ratio of 93 nucleotide to 61 nucleotide fragments increased accordingly (lane 10). Thus transcription of the rDNA was not detectably inhibited by high levels of C-G methylation. The experiment also argues strongly against the possibility that a small fraction of sperm rDNA is incompletely methylated and is responsible for all the transcription of sperm rDNA that we

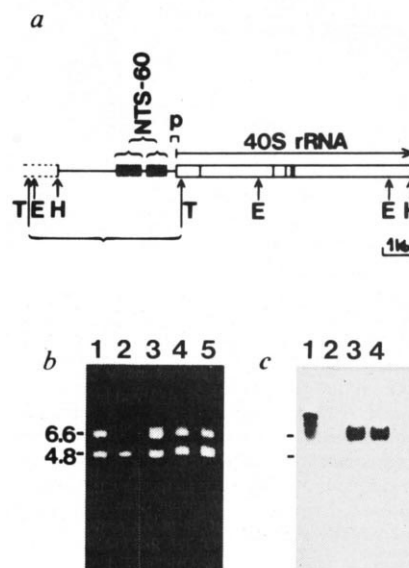
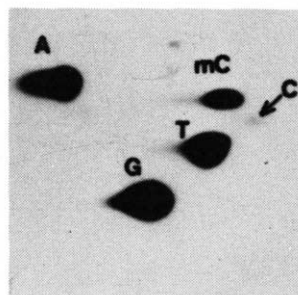


Fig. 1 Structure and methylation of blood and sperm rDNA of *X. laevis*. **a**, A map of the rDNA repeat unit showing sites for *Eco*RI (E) and *Hind*III (H). Two *Taq*I sites flanking the spacer are shown (T). Other *Taq*I sites (not shown) are in the transcribed region. One *Taq*I site is located 93 bp downstream of the 40S transcription start site and this site was labelled for sequencing studies and as a probe for S₁-mapping experiments. The promoter, p, lies between -146 bp and +16 bp¹⁶. Typically, two clusters of NTS-60 repeats are present, though up to eight clusters have been observed. **b**, Absence of undermethylated sites in purified sperm rDNA. The figure shows an ethidium bromide stained agarose gel of: lane 1, *Eco*RI-digested blood DNA; lane 2, *Eco*RI plus *Hpa*II-digested blood DNA; lane 3, *Eco*RI-digested sperm DNA; lane 4, *Eco*RI plus *Hpa*II-digested sperm DNA; lane 5, *Eco*RI plus *Hpa*II plus *Ava*I-digested sperm DNA. **c**, Retention of methylated C-G in sperm rDNA after injection into *X. borealis* oocytes. The figure shows a Southern blot, hybridized to a probe specific for the *X. laevis* rDNA spacer, of the following: lane 1, undigested DNA from injected oocytes; lane 2, DNA from uninjected oocytes; lane 3, DNA from injected oocytes digested with *Eco*RI; lane 4, DNA from injected oocytes digested with *Eco*RI plus *Hpa*II.

Methods: For **b**, ribosomal DNA was purified from blood or from sperm heads of 5–20 *Xenopus* males. Preparation of sperm heads by sonication of a crude nuclear suspension from total testis DNA has been described²⁵. Total DNA was sheared in a Dounce homogeniser, and centrifuged to equilibrium in a solution containing 0.14 mg ml⁻¹ DNA, 0.78 g ml⁻¹ CsCl, 60 μ g ml⁻¹ actinomycin D, 10 mM Tris, 1 mM EDTA (pH 8.5). Centrifugation was at 26,000 r.p.m. for 3 days at 3 °C in a Sorvall vertical titanium rotor (34 ml per tube). The rDNA peak was localized on the light side of the main peak by ethidium bromide staining of aliquots. The peak was then pooled, precipitated with 3 vols 70% ethanol, and redissolved in 4.2 ml of CsCl solution (density 1.710 g ml⁻¹). This solution was centrifuged at 45,000 r.p.m. (18 °C) for 20 h in a Sorvall vertical rotor (4.2 ml per tube). Purified rDNA fractions were pooled and ethanol precipitated. The yield was usually 1–1.5 μ g per mg total DNA. About 0.1 μ g of blood (lanes 1 and 2), or sperm (lanes 3–5) rDNA was restriction enzyme-digested and analysed by agarose gel electrophoresis and ethidium bromide staining. In blood rDNA the spacer (upper) fragment is cut at unmethylated sites in the clusters of NTS-60⁵. For **c**, 2 μ g of sperm rDNA was digested with *Eco*RI, diluted to 5 μ g ml⁻¹ and ligated overnight in standard conditions. Electron microscopy and agarose gel electrophoresis showed that over 50% of rDNA fragments had formed circles. The DNA was phenol extracted, ethanol-precipitated, and dissolved in MMR²⁶ at about 20 μ g ml⁻¹. About 10 nl of this solution were injected into centrifuged eggs²⁷. After 4 h nucleic acids were extracted and analysed on an agarose gel by blotting¹² and hybridizing to a probe that is specific for the *X. laevis* spacer. As a control for complete digestion, 0.1 μ g of lambda DNA was included in the digests and the fragment pattern was checked by ethidium bromide staining.

Fig. 2 High levels of 5-methylcytosine in the C-G sequences of sperm rDNA. Sperm rDNA (0.1 μ g) was nicked by sonication or DNaseI and labelled with DNA polymeraseI and [α - 32 P] dGTP. Cold triphosphates other than dGTP were not present⁹. Nearest neighbour analysis was as described^{9,28}. Scintillation counting the spots gave unmethylated C-G as about 3% of total C-G. This value may be accurate for sperm rDNA, or there may be contaminating sequences in the rDNA that contribute to unmethylated C-G. We have reservations about the method, however, because total C levels in sperm rDNA were consistently under-estimated by over 50%. This was not found when cloned rDNA was analysed or when the other deoxynucleoside triphosphates were present during labelling (D. Macleod, unpublished data), suggesting that 5-methyl C itself is discriminated against. The value of 97% methylation of C-G includes a correction for this.



observe. If this were the case, equal amounts of sperm and cloned circles would not be expected to be transcribed equally.

The 40S promoter of *X. laevis* rDNA has been localized between positions -145 and +16, as deletions within this region eliminate transcription in the oocyte¹⁰. In the light of our results with sperm rDNA, we wished to test directly the level of methylation at every cytosine residue in the promoter. We labelled the *Taq*I site at +93 bp on either the 5' or 3' strand, and sequenced upstream using the chemical cleavage method¹⁷ (Fig. 4). As methyl C residues are not cleaved by any of the chemical reactions, their presence is indicated by missing nucleotides in the sequence ladder that are opposite G on the other strand. For comparison, the promoter of pXlr108c, which has been sequenced previously¹⁸, was cleaved at C only and run next to the sperm rDNA ladder. Surprisingly, the promoter sequence of sperm rDNA (pooled from 20 animals, each with about 1000 ribosomal genes per cell) was homogeneous and identical to the sequence of pXlr108c. Moreover, of the 19 C-Gs in the promoter sequence, all were methylated on both strands in sperm rDNA (Fig. 4). The missing C residues only occurred at C-G, and there was no evidence for a minor fraction of unmethylated C at these sites.

Our results show that complete methylation of C-G in the promoter and heavy methylation of C-Gs in the rest of the spacer does not interfere with rDNA transcription in oocytes. Similar conclusions were reached when 5S DNA from blood cells was tested in oocytes¹⁹. The 5S genes were at least partially methylated at most C-Gs in the repeat unit^{20,21}, but were transcribed efficiently. Transcription in the oocyte of two viral genes, however, is inhibited by *in vitro* methylation at *Hpa*II sites^{13,14}. There is also good evidence that infectivity and transcription of Moloney murine leukaemia virus²², and expression of the hamster adenosyl phosphoribosyl transferase gene in cultured cells²³ is inhibited by C-G methylation.

The simplest explanation of these results is that DNA methylation is irrelevant to rDNA expression *in vivo* and *in vitro*. If correct, this would suggest that rDNA differs from those RNA polymerase II genes whose transcription is inhibited by C-G methylation. There are, however, two alternative explanations of the data that must be considered. The first alternative explanation is that the oocyte does not respond correctly to methylation and is therefore of limited use as an assay system. Against this argument, it should be recalled that methylation of two viral genes has been shown to inhibit their transcription in oocytes^{13,14}. The second alternative explanation of the data is that the oocyte may be insensitive to upstream regions of the spacer at which changes in methylation take place. In the simplest experiments, much of the spacer, including the under-methylated regions, can indeed be deleted without affecting transcription in oocytes. Embryos, on the other hand, require spacer sequences for correct expression of injected rDNA²⁴.

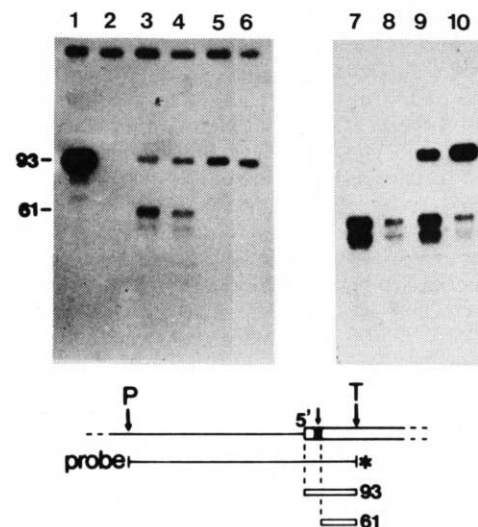


Fig. 3 Transcription of *X. laevis* sperm rDNA in *X. borealis* oocytes. The figure shows the results of *S*₁-nuclease protection analyses, using as a probe fragment P-T as shown in the diagram below the autoradiogram, of RNA from: lane 1, *X. laevis* oocytes; lane 2, *X. borealis* oocytes; lanes 3-10, *X. borealis* oocytes injected with: lane 3, pX1ΔES1 (see below) plus circularized sperm rDNA *Eco*RI fragments (20 μ g ml⁻¹); lane 4, pX1ΔES1 plus circularized blood rDNA *Eco*RI fragments (20 μ g ml⁻¹); lane 5, circularized sperm rDNA alone; lane 6, circularized blood rDNA alone; lane 7, pX1ΔES1 (5 μ g ml⁻¹) alone; lane 8, pX1ΔES1 (5 μ g ml⁻¹) plus linear (uncut) sperm rDNA; lane 9, pX1ΔES1 (5 μ g ml⁻¹) plus circularized sperm rDNA (5 μ g ml⁻¹); lane 10, pX1ΔES1 (5 μ g ml⁻¹) plus circularized sperm rDNA (20 μ g ml⁻¹). The sperm rDNA tested in lanes 9 and 10 was digested with *Hpa*II before injection to linearize any contaminating unmethylated rDNA. Linear rDNA is transcribed minimally in this system (see lane 8).

Methods: *Eco*RI fragments of sperm rDNA were circularized and injected into centrifuged oocytes (see legend to Fig. 1). The concentration of circularized large *Eco*RI fragments was estimated from ethidium bromide-stained gels of the ligated rDNA. After 4-18 h incubation at 20-25 °C, total RNA was extracted²⁷ and subjected to *S*₁-nuclease protection analysis¹⁵. The probe was a *Pst*I-*Taq*I fragment of pXlr108c (see fragment P-T, in the diagram), which was labelled with [γ - 32 P]ATP and polynucleotide kinase, and strand separated on a polyacrylamide gel (40:1 acrylamide:bis-acrylamide). The single-stranded probe was annealed with RNA from five injected oocytes overnight at 65 °C in 0.9M NaCl, 0.09M Na citrate, 50% formamide²⁹. Annealing mixes were treated with *S*₁-nuclease, precipitated with ethanol, and analysed on an 8% polyacrylamide-7M urea gel. End-labelled pAT153 *Hpa*II fragments were included on each gel as size markers. Plasmid pX1ΔES1 (a gift from Tom Moss, Portsmouth Polytechnic) is the large *Eco*RI fragment of rDNA inserted into a modified pBR322 vector, but with a deletion in the transcribed region (shaded area arrowed in diagram). The transcript from this plasmid therefore protects a 61 nucleotide fragment from *S*₁-nuclease digestion instead of the full-sized 93 nucleotide fragment¹⁰. A second band at 52 nucleotides arises from an AT-rich sequence downstream of the deletion (see refs 10 and 16).

However, if the deleted genes are co-injected into oocytes with undeleted rDNA as a competitor, the deleted rDNA is transcribed very inefficiently¹⁶. Thus oocytes become sensitive to upstream spacer sequences in a competition experiment with unmethylated rDNA. As we showed that methylated rDNA is transcribed efficiently in a competition experiment, it is clear that methylation of upstream spacer sequences is not equivalent to their deletion.

If loss of methylation is not a prerequisite for rDNA transcription, how can the undermethylation of a specific, evolutionarily-conserved⁵ sequence in the spacer be explained? Previous work has shown that the pattern of unmethylated C-Gs within a copy of NTS-60 is compatible with the binding of a protein which interferes to varying degrees with the methylation of C-Gs within the sequence⁵. While it is conceivable that the function

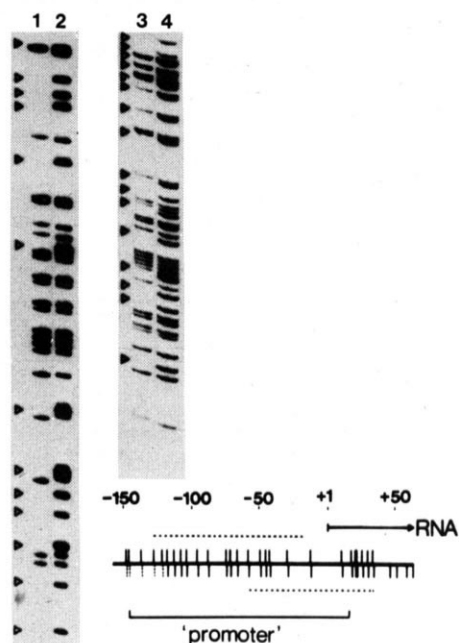


Fig. 4 Methylation at all C-Gs in the promoter of *X. laevis* rDNA. The figure shows examples of Maxam-Gilbert sequencing gels of sperm (lanes 1 and 3) and cloned (pXlr108c, lanes 2 and 4) rDNA fragments after partial cleavage at C residues. Missing bands in the sperm track (arrow heads) indicate 5-methyl C residues. Complete sequencing of the sperm rDNA promoter region showed that each missing band is opposite a G on the other strand, and is part of the sequence 5' C-G. The regions covered in the examples are from overlapping segments on opposite strands as shown by the lower (lanes 1 and 2) and upper (lanes 3 and 4) broken lines on the diagram. The diagram shows the promoter region (bracket) and the position of 5-methyl C residues on each strand (vertical lines). All the C-Gs that were encountered on both strands were methylated. Methyl C residues marked by broken lines are in a region where the sequence could not be read properly on this strand. Their positions were inferred from missing C bands in sperm tracks, and from the sequence on the other strand.

Methods: 2 µg of sperm rDNA was digested with *TaqI* to give several small fragments from within the gene, plus a large fragment which includes the whole spacer and 93 bp from the 5' end of the sequence coding for 40S pre-rRNA (see Fig. 1a). The fragments were treated with calf intestinal alkaline phosphatase, phenol and chloroform extracted, and precipitated. Ends were labelled with [γ - 32 P]ATP and polynucleotide kinase, or [α - 32 P]dCTP and the Klenow fragment of DNA polymerase I. Labelled DNA was purified on a Sephadex G-75 column and digested with *HindIII* and *HpaII* to remove the upstream labelled end and to remove any contaminating blood rDNA. The mixture was then applied to a 1.2% low melting point agarose gel and the heterogeneous group of spacer bands at about 4.5 kb was eluted. The equivalent fragment of pXlr108c was purified by the same procedure but omitting *HpaII* digestion. Sequencing was by the Maxam-Gilbert technique¹⁷ which does not cleave at 5-methyl C residues.

of this putative protein binding may be to demethylate a copy of NTS-60, it is also possible that demethylation is an accidental consequence of an interaction whose real function lies elsewhere.

We thank Alan Coleman, Aimée Bakken, Tom Moss, Ed Southern, Brian McStay and Anne Deane for their help. The work was supported by the MRC.

Received 7 July; accepted 24 August 1983.

- Lewin, B. *Gene Expression* Vol. 2 p. 865 (Wiley, New York, 1980).
- Long, E. & Dawid, I. B. *A. Rev. Biochem.* **49**, 727-764 (1980).
- Bird, A. P. & Southern, E. M. *J. molec. Biol.* **118**, 27-47 (1978).
- Bird, A. P., Taggart, M. H. & Macleod, D. *Cell* **26**, 381-390 (1981).
- La Volpe, A., Taggart, M. H., Macleod, D. & Bird, A. P. *Cold Spring Harb. Symp. quant. Biol.* **47**, 585-592 (1982).
- Boseley, P. G., Moss, T., Machler, M., Portmann, R. & Birnstiel, M. L. *Cell* **17**, 19-31 (1979).
- Sollner-Webb, B. & Reeder, R. H. *Cell* **18**, 485-499 (1979).
- Birnstiel, M. L., Wallace, H., Sirlin, J. L. & Fischberg, M. *Natn. Cancer Inst. Monogr.* **23**, 431-447 (1966).

- Gruenbaum, Y., Stein, R., Cedar, H. & Razin, A. *FEBS Lett.* **124**, 67-71 (1981).
- Moss, T. *Cell* **30**, 835-842 (1982).
- Sollner-Webb, B. & McKnight, S. L. *Nucleic Acids Res.* **10**, 3391-3405 (1982).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Vardimon, L., Kressmann, A., Cedar, H., Maechler, M. & Doerfler, W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1073-1077 (1982).
- Fradin, A., Manley, J. L. & Prives, C. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5142-5146 (1982).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 721-732 (1977).
- Moss, T. *Nature* **302**, 223-228 (1983).
- Maxam, A. M. & Gilbert, W. *Meths Enzymol.* **65**, 499-560 (1980).
- Moss, T., Boseley, P. G. & Birnstiel, M. L. *Nucleic Acids Res.* **8**, 467-485 (1980).
- Brown, D. D. & Gurdon, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2064-2068 (1977).
- Fedoroff, N. V. & Brown, D. D. *Cell* **13**, 701-716 (1978).
- Miller, J. R., Cartwright, E. M., Brownlee, G. G., Fedoroff, N. V. & Brown, D. D. *Cell* **13**, 717-725 (1978).
- Stuhlmann, H., Jahner, D. & Jaenisch, R. *Cell* **26**, 221-232 (1981).
- Cedar, H. *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 605-609 (1983).
- Busby, S. & Reeder, R. H. (1983) *Cell* (in the press).
- Bird, A. P., Taggart, M. H. & Gehring, C. J. *molec. Biol.* **152**, 1-17 (1981).
- Newport, J. & Kirschner, M. *Cell* **30**, 675-686 (1982).
- Kressman, A., Clarkson, S. G. & Telford, J. L. *Cold Spring Harb. Symp. quant. Biol.* **42**, 1077-1082 (1978).
- Dawid, I. B., Brown, D. D. & Reeder, R. H. *J. molec. Biol.* **51**, 341-360 (1970).
- Birnstiel, M. L., Sells, B. H. & Purdom, I. F. *J. molec. Biol.* **63**, 21-39 (1972).

Evolutionary adaptation of plasmid-encoded enzymes for degrading nylon oligomers

Hirosuke Okada, Seiji Negoro, Hiroyuki Kimura & Shunichi Nakamura

Department of Fermentation Technology, Osaka University, Yamada-oka, Suita-shi, Osaka 565, Japan

Flavobacterium sp. KI72 metabolizes 6-aminohexanoic acid cyclic dimer, a by-product of nylon manufacture¹, through two newly evolved enzymes, 6-aminohexanoic acid cyclic dimer hydrolase (EI)² and 6-aminohexanoic acid linear oligomer hydrolase (EII)³. These enzymes are active towards man-made compounds, the cyclic dimer and linear oligomers of 6-aminohexanoic acid respectively, but not towards any of the natural amide bonds tested^{2,3}. The structural genes of EI (*nylA*) and EII (*nylB*) are encoded on pOAD2, one of three plasmids harboured in *Flavobacterium* sp. KI72^{4,5}. This plasmid contains two kinds of repeated sequence (RS-I and RS-II); one of the two RS-II sequences, RS-II_A, contains the *nylB* gene⁶, while the other, RS-II_B, contains a homologous *nylB'* gene. From comparisons of the nucleotide sequences and gene products of the *nylB* and *nylB'* genes, we now conclude that EII enzyme is newly evolved by gene duplication followed by base substitutions on the same plasmid.

Figure 1 is a physical and functional map of pOAD2 showing the loci of the *nylA* and *nylB* genes and the RS-I and RS-II sequences⁶. Figure 2 compares the restriction sites on RS-II_A and RS-II_B, of which those marked by triangles are conserved in both sequences. These results indicate that the sequences are homologous but not identical.

Using Maxam and Gilbert's method⁷, we determined the DNA base sequence of RS-II_A in pNL21 (consisting of pBR322⁸, a 207 base pair (bp) *EcoRI* fragment of pKB252⁹ encoding the *lacUV5* promoter and the 2.7-5.0 kilobase pair (kbp) region on the pOAD2 map) and RS-II_B in pNDH10 (consisting of pBR322 and the 13.5-16.8-kbp region on the pOAD2 map). The strategy of sequencing is shown in Fig. 2.

Figure 3 shows the result of the DNA sequencing. In the RS-II_A region, the only open frame available to encode a peptide of 42,000 molecular weight starts from ATG, the initiation codon, which has a Shine-Dalgarno sequence¹⁰ of GGAG¹¹. This open frame has a termination codon TAG 1,176 bp after the initiation codon, and would encode a peptide of 392 amino acids. This size agrees well with the molecular weight of the subunit of EII enzyme (42,000)³.

EII enzyme produced by *Escherichia coli* harbouring pNL3 (consisting of pBR322, a 207-bp *EcoRI* fragment of pKB252 encoding the *lacUV5* promoter and the 2.1-4.1-kbp region on

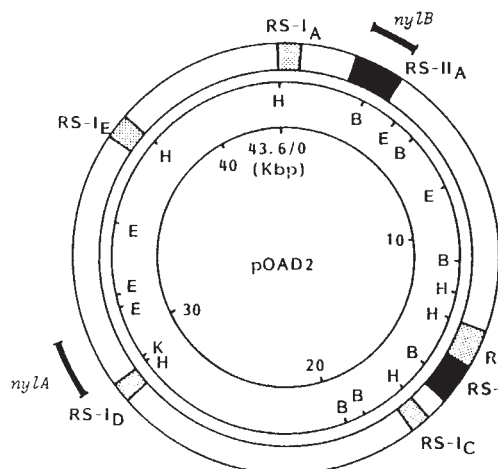


Fig. 1 Physical and functional map of pOAD2. H, B, E and K indicate *Hind*III, *Bgl*II, *Eco*RI and *Kpn*I sites, respectively. *nylA* (28.8–30.8-kbp region) and *nylB* (2.7–4.1-kbp region) genes are indicated. One type of repeated sequence (RS-I) (shown as a shaded bar) is located in five regions, 0–0.77 kbp (RS-I_A), 13.5–14.7 kbp (RS-I_B), 16.8–17.4 kbp (RS-I_C), 28.2–28.8 kbp (RS-I_D) and 37.1–37.9 kbp (RS-I_E). Another type of repeated sequence (RS-II) (shown as a solid bar) is located in two regions, 2.7–4.2 kbp (RS-II_A) and 14.7–16.1 kbp (RS-II_B).

the pOAD2 map) was purified and subjected to amino acid analysis. The numbers of each amino acid residue contained in the EII enzyme subunit were estimated from the results of amino acid analysis (numbers in parentheses) and from the DNA sequence: Lys,3(3.0); His,7(7.0); Trp,11(8.1); Arg,31(30.8); Asp+Asn,41(40.7); Thr,28(26.5); Ser,26(23.9); Glu+Gln,36(36.6); Pro,21(20.0); Gly,39(38.2); Ala,39(38.9); Cys,5(1.1); Val,32(32.4); Met,6(5.2); Ile,10(9.9); Leu,36(36.3); Tyr,12(11.8); Phe,9(9.0). These two estimates agreed for all amino acids except tryptophan and cysteine, which are unstable. The amino acid sequence of the N-terminal of this enzyme was Met-Asn-Ala... contaminated by a minor portion (~30%) of Asn-Ala-Arg..., which might be formed by elimination of the terminal methionine residue of the major protein (unpublished results). This amino acid sequence indicates that EII protein is not fused to a part of β -galactosidase.

The DNA sequence of RS-II_B also contains an open frame able to encode a peptide of 392 amino acids (EII'). This frame forms part of a homologous region with RS-II_A covering *nylB*. The *nylB* gene and the EII' structural gene (*nylB'*) show 88% homology. The base-pair substitutions in the frame are not distributed evenly, with 83 of the 140 occurring at the third base of the codon, of which 69 do not change the encoded amino acids. From the base sequences of the *nylB* and *nylB'* genes, their gene products should have 345 identical amino acids out of 392 (88% homology).

pNDH10, which encodes the *nylB'* gene, cannot produce EII' protein detectable by anti-EII serum after transformation of *E. coli* 294 strain⁹. To produce EII' protein in *E. coli*, we inserted the 207-bp *Eco*RI fragment containing *lacUV5* promoter (after filling up the cohesive ends)¹² at the *Sma*I site located 218 bp upstream from the initiation codon of the *nylB'* gene in pNDH10 (the new hybrid plasmid was named pNDH10L1). In a double diffusion experiment¹³, purified EII protein and the cell lysate of *E. coli* harbouring pNDH10L1 produced precipitin lines to anti-EII serum, which fused at the corner with a spur, while *E. coli* (pNDH10) cell extract formed no precipitin line. This result shows the immunological homology of the EII and EII' proteins. The EII and EII' proteins had the same electrophoretic mobility on SDS-polyacrylamide gel detected by anti-EII serum and ¹²⁵I-protein A after transfer to a nitrocellulose filter, indicating that they have the same molecular weight.

Figure 4 shows the results of quantitative immunoelectrophoresis using anti-EII serum. A linear relationship was

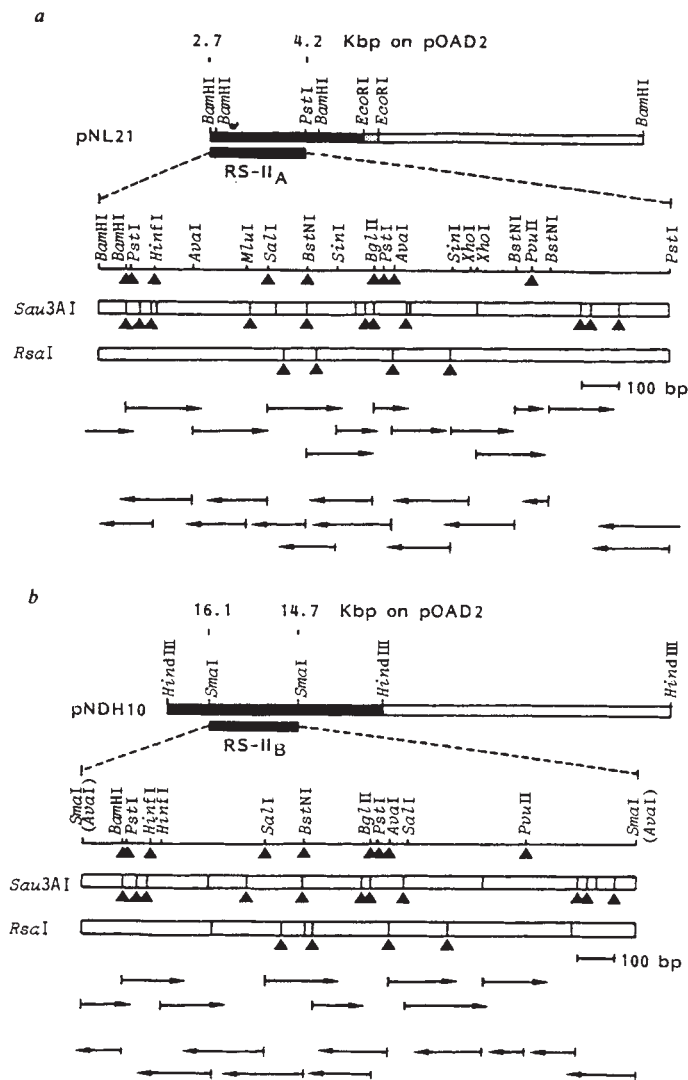
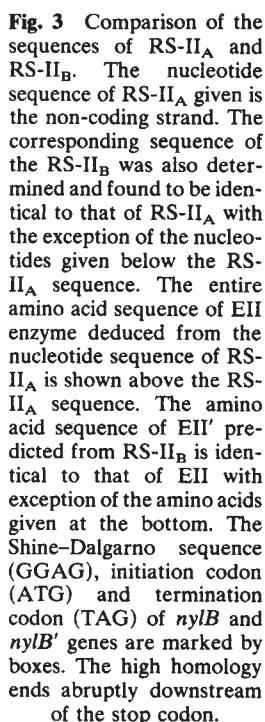


Fig. 2 Restriction map and sequencing strategy of RS-II_A (a) and RS-II_B (b). The shaded bar indicates a 207-bp *Eco*RI fragment encoding the *lacUV5* promoter.

Methods: Restriction sites (*Ava*I, *Bam*HI, *Bgl*II, *Bst*NI, *Hinf*I, *Mlu*I, *Pst*I, *Sal*I, *Sin*I, *Sma*I and *Xho*I sites) were determined previously⁶. For determination of *Sau*3AI and *Rsa*I sites on RS-II_B, 800-bp *Ava*I and 650-bp *Ava*I fragments of pNDH10, which encode RS-II_B, were end-labelled with [γ -³²P]ATP⁷, then the former was digested with *Sal*I, the latter with *Pvu*II, followed by electrophoresis. The purified *Ava*I-*Sal*I (480 bp), *Ava*I-*Sal*I (320 bp), *Ava*I-*Pvu*II (360 bp) and *Ava*I-*Pvu*II (290 bp) fragments were digested with *Sau*3AI or *Rsa*I in limited conditions, and the cleavage sites were deduced according to the method of Smith and Birnstiel¹⁵. The restriction sites of RS-II_A were similarly determined. The nucleotide sequences were determined by the chemical method of Maxam and Gilbert⁷. Labelling of the 5' ends of DNA fragments with [γ -³²P]ATP and T4 polynucleotide kinase is described in ref. 7. In some experiments, the 3' ends of the DNA fragments were labelled with [α -³²P]cordycepin-5'-triphosphate using terminal deoxynucleotidyl transferase¹⁶. Sequencing polyacrylamide-urea gels were made and run as described in ref. 7. ³²P-labelled ends are represented by short vertical lines, and the extent and direction of sequencing are indicated by horizontal arrows. Restriction sites conserved in RS-II_A and RS-II_B are shown by triangles.

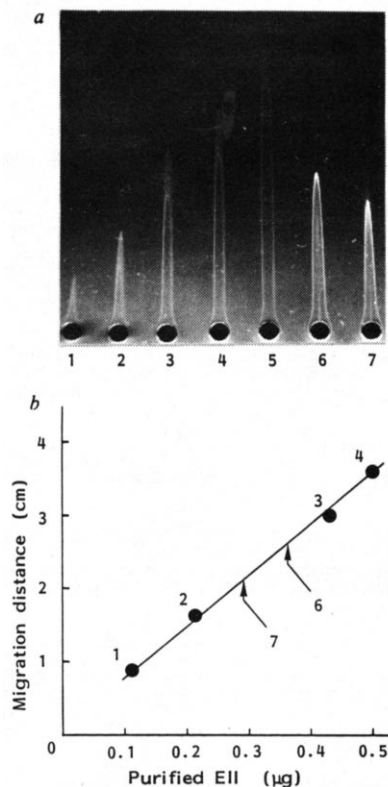
observed between the amount of purified EII enzyme (<0.5 μ g) in the well and the peak height in the agarose plate. Because of possible differences in antigenicity between EII and EII' proteins, however, quantitative antigenicity may not directly reflect the amount of protein. The enzyme activity was measured by the decrease in 6-aminohexanoic acid linear dimer. A 10-fold greater concentration of cell lysate from *E. coli* 294



We thank F. Sakiyama and S. Tsunazawa for the amino acid analysis and the partial amino acid sequencing of EII enzyme, K. Ushiyama and T. Kawasaki for the preparation of the anti-EII serum, and N. Kanzaki for technical help. This work was partly

Fig. 4 a, Quantitative immunoelectrophoresis of EII and EII' proteins. Slots 1–5, purified EII; 0.11 μ g (slot 1), 0.21 μ g (slot 2), 0.43 μ g (slot 3), 0.50 μ g (slot 4), 0.85 μ g (slot 5); slot 6, cell extract of *E. coli* 294 harbouring pNDH10L1 (EII'-producing) (32 μ g); slot 7, cell extract of *E. coli* 294 harbouring pNL3 (EII-producing) (30 μ g). **b**, Determination of the EII and EII' content in the cell extract of *E. coli* 294. A linear relationship was observed between the amounts of EII (slots 1–4) and the peak height of the immunoprecipitate. The peak heights of the precipitates obtained from *E. coli* (pNDH10L1) (slot 6) and *E. coli* (pNL3) (slot 7) are indicated by arrows.

Methods: Rocket quantitative immunoelectrophoresis was carried out in 15 ml of 1.2% agarose gel containing 10% anti-EII serum on a glass plate (8.5 $\times$ 10.5 cm) according to the procedure of Laurell¹⁷. 5 μ l of each sample was applied to a gel slot and electrophoresis was carried out at 20 mA for 10 h.



supported by Grants-in-Aid for Scientific Research 56030023 and 57220014 from the Ministry of Education, Culture, and Science, Japan.

Received 13 June; accepted 7 September 1983.

- Kinoshita, S., Kageyama, S., Iba, K., Yamada, Y. & Okada, H. *Agric. biol. Chem.* **39**, 1219–1223 (1975).
- Kinoshita, S. *et al. Eur. J. Biochem.* **80**, 489–495 (1977).
- Kinoshita, S. *et al. Eur. J. Biochem.* **116**, 547–551 (1981).
- Negoro, S. *et al. J. Bact.* **143**, 238–245 (1980).
- Negoro, S. & Okada, H. *Agric. biol. Chem.* **46**, 745–750 (1982).
- Negoro, S., Taniguchi, T., Kanaoka, M., Kimura, H. & Okada, H. *J. Bact.* **155**, 22–31 (1983).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
- Bolivar, F. *et al. Gene* **2**, 95–113 (1977).
- Backman, K., Ptashne, M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4174–4178 (1976).
- Shine, J. & Dalgarno, L. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1342–1346 (1974).
- Gold, L. *et al. A. Rev. Microbiol.* **35**, 365–403 (1981).
- Jacobsen, H., Klenow, H. & Overgaard-Hansen, K. *Eur. J. Biochem.* **45**, 623–627 (1974).
- Ouchterlony, O. *Acta path. microbiol.* **26**, 507–515 (1949).
- Paterson, A. & Clarke, P. H. *J. gen. Microbiol.* **114**, 75–85 (1979).
- Smith, H. O. & Birnstiel, M. L. *Nucleic Acids Res.* **3**, 2387–2398 (1976).
- Tu, C. D. & Cohen, S. N. *Gene* **10**, 177–183 (1980).
- Laurell, C. B. *Analyt. Biochem.* **15**, 45–52 (1966).

Molecular identification of a human DNA repair gene following DNA-mediated gene transfer

Jaime S. Rubin, Alexandra L. Joyner,
Alan Bernstein & Gordon F. Whitmore

Department of Medical Biophysics, University of Toronto and
Physics and Bioresearch Divisions, Ontario Cancer Institute,
Toronto, Ontario, Canada M4X 1K9

Although it has long been evident that the response of eukaryotes to DNA damaging agents is determined by the effectiveness of a variety of DNA repair systems, there is little detailed knowledge of the nature of these systems or the genes which control them. In humans, a number of hereditary conditions, including xeroderma pigmentosum, ataxia telangiectasia and Fanconi's anaemia, exhibit increased sensitivity to a variety of DNA damaging agents and a predisposition to cancer, suggest-

ing a defect in some aspect of DNA repair¹. This report describes the identification of a human DNA repair gene following DNA-mediated gene transfer into Chinese hamster ovary (CHO) mutant cells^{2,3}, that like xeroderma pigmentosum cells, are sensitive to a variety of DNA damaging agents and are defective in the initial incision step of DNA repair⁴. The resulting transformants exhibit normal resistance to DNA damaging agents and independent transformants demonstrate a common set of human DNA sequences associated with a human DNA repair gene. These observations provide the basis for the isolation and characterization of the human genes responsible for DNA repair.

The CHO recipient cells (UV-20 and UV-41) chosen for gene transfer represent two complementation groups⁵ of DNA repair deficient mutants. They were selected for their sensitivity to ultraviolet (UV) light and possess several other characteristics useful in transfection experiments: (1) in contrast to xeroderma pigmentosum cells⁶, the CHO mutants are extremely sensitive to mitomycin-C (MM-C)², which allows efficient selection of resistant transformants under conditions of chronic exposure; (2) they are capable of stably incorporating and expressing foreign DNA at a relatively high efficiency; (3) the transfection of human donor DNA into CHO recipient cells allows the use of highly repetitive human DNA sequences (*Alu* sequences) as probes to verify DNA transfer^{7–9} and to provide an initial characterization of the exogenous gene.

To reduce the possibility of selecting a DNA repair proficient CHO revertant, use was made of a co-selection procedure requiring a selected cell to express both repair proficiency and a drug resistance phenotype, conferred by the expression of a transfected cloned bacterial gene. This approach is based on the observation that selection for the expression of a cloned gene identifies a small subpopulation of cells that are competent for the stable uptake of exogenous DNA sequences¹⁰. In the presence of an excess of a cloned gene, cells transformed for this marker have a high probability of transformation by a second marker, so the frequency of co-transformation is much higher than the frequency of transformation for one marker along with spontaneous mutation for the second. In the selection protocol used, recipient cells were exposed to both high molecular weight human (HeLa) DNA and the plasmid pSV2-GPT containing the dominant *ecogpt* gene from the bacterium *Escherichia coli*¹¹. After calcium phosphate mediated gene transfer^{12,13}, cells were initially selected for the expression of the *ecogpt* gene in growth medium containing mycophenolic acid (MAX), followed by growth in MM-C to select for repair proficiency (see Fig. 1 legend for experimental details).

On the basis of the number of cells initially plated prior to DNA exposure, the frequency of MAX resistant transfectants varied between 3×10^{-4} and 2×10^{-3} , and the co-transfer frequency (both MAX and MM-C resistance) in successful experiments varied between 6×10^{-8} and 2×10^{-7} (see Fig. 1 legend).

Two primary transformants, X25 and X38, were selected from UV-20 (Fig. 1) and a secondary transformant, X25-37, was generated by transferring DNA from X25 along with an excess of the pSV2-GPT plasmid. Similarly, one primary, R35, and one corresponding secondary transformant, R35-26, were generated from UV-41. Prior to further testing, the transformants were grown continuously in culture media containing MM-C at the concentration used for selection. As a test of genetic stability, the transformants were grown for 6 weeks in non-selective medium and all retained their resistance to MM-C, indicating stable expression of the DNA repair function. After many months of continuous exposure to MM-C, but in the absence of selection for the *ecogpt* gene, the transformants retained at least a 70% plating efficiency in medium containing MAX, indicating stable integration of the *ecogpt* gene.

To verify that the transformants had wild-type resistance to both UV and MM-C, survival curves were determined for all cell lines used in these experiments (Fig. 2). As shown, both mutant lines were sensitive to UV light (7-fold) and extremely

Fig. 1 Generation of DNA repair proficient transformants.

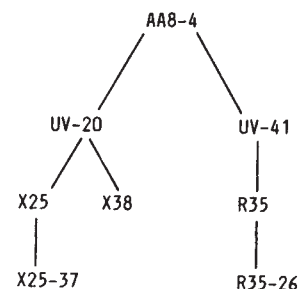
Methods: 24 h before DNA transfection, 5×10^5 recipient CHO cells were added to a 100 mm Petri dish in 10 ml of growth medium (α medium¹⁶ without nucleotides supplemented with antibiotics and 10% fetal calf serum). 20 μ g of high molecular weight DNA, extracted from human HeLa cells, was calcium-precipitated^{12,13} with 5 μ g of the plasmid pSV2-GPT¹¹. High molecular weight DNA was isolated by incubating proteinase K (200 μ g ml⁻¹) and SDS (0.2%) with cells suspended in buffer A (10 mM Tris HCl, pH 7.9, 2 mM EDTA and 0.4 mM NaCl) for 2 h at 37 °C. After extraction with buffered phenol, phenol/chloroform (50:50) and phenol/chloroform/isoamyl alcohol (49:49:2), the DNA was ethanol precipitated, spooled on a glass rod and resuspended in TE buffer (1 M Tris-HCl, pH 7.8, 0.1 mM EDTA). The cells were incubated with the calcium phosphate-DNA precipitate for 24 h, after which the growth medium was changed. One day later, the dishes that now contained $\sim 4.4 \times 10^6$ viable cells were trypsinized and replated in six 100 mm Petri dishes containing 10 ml of MAX medium (growth medium plus 20 μ g ml⁻¹ of mycophenolic acid (Eli Lilly and Co.), 25 μ g ml⁻¹ of adenine and 250 μ g ml⁻¹ of xanthine). Four to five days later, the media were changed. This procedure typically yielded 20–200 MAX resistant colonies on each of the six plates. When MAX resistant colonies contained 10–100 cells, 1 ml of 10^{-7} M MM-C (Bristol Laboratories) was added to each plate. In a typical experiment of 40 initial DNA transfection dishes, up to two MM-C resistant colonies were detected 3–6 weeks following trypsinization. These colonies were cloned and maintained in culture in growth medium containing nucleotides and 10^{-8} M MM-C. To generate secondary transformants, high molecular weight DNA extracted as described above from primary transformants along with 5 μ g of pSV2-GPT was used to generate secondary transformants in another round of DNA transfection.

CHO Wild type

DNA repair deficient mutant

1' Transformant

2' Transformant



sensitive to MM-C (70-fold) when compared with the wild-type CHO parent, AA8-4. It is also apparent from this figure that the primary and secondary transformants are resistant to both agents, providing further evidence that these cells are now phenotypically DNA repair proficient. Figure 2 also demonstrates that the survival curves of the transformants are not identical to either each other, the CHO wild-type (AA8-4), or the DNA donor (HeLa). This observation is consistent with the understanding that other factors, including levels of expression of the transfected gene and the contributions of other cellular repair enzymes determine the final phenotype of these cell lines.

In addition to the biological evidence presented above, molecular evidence is required to establish that the repair proficient phenotype of the transformants is due to the integration

of transfected DNA. Gel electrophoresis combined with blot hybridization¹⁴ was used for this purpose. As indicated in Fig. 3, all five transformants contained multiple copies of the plasmid pSV2-GPT indicating that the recipient cell lines were competent for the uptake and stable integration of exogenous DNA.

To demonstrate the successful integration of human DNA, use was made of the fact that the human genome contains a family of highly repetitive 'Alu' sequences, which, under stringent hybridization conditions, can be used to establish the presence of human DNA⁷⁻⁹. This ability to distinguish human DNA against a CHO background provides a dimension to this study not available in a previous report of the transfer of a human DNA repair gene¹⁵. Cellular DNA from the primary and secondary transformants was analysed for the presence of human

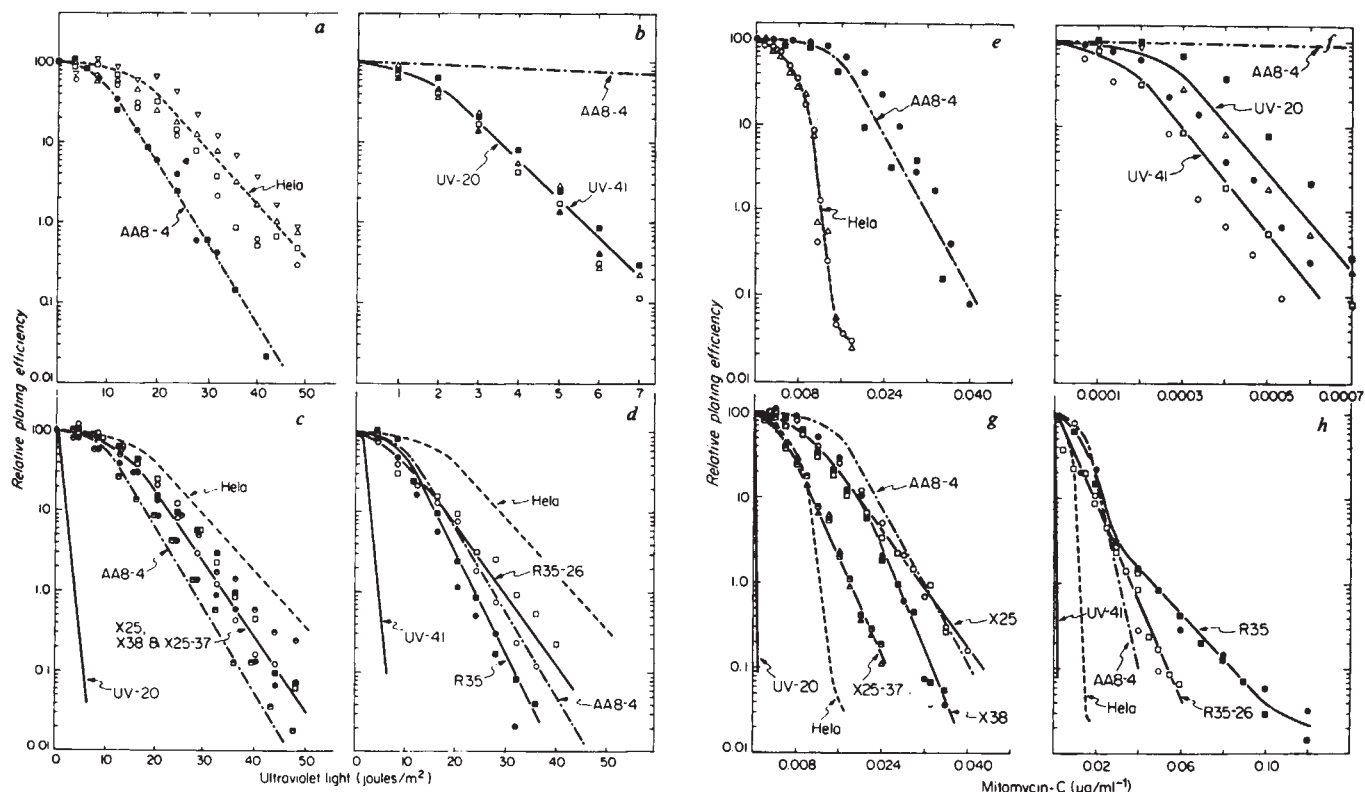


Fig. 2 Survival curves of cell lines following exposure to UV light (a–d) and MM-C (e–h). a, e: AA8-4, closed symbols; HeLa, open symbols; b, f: UV-20, closed symbols; UV-41, open symbols; c, g: X25, open symbols; X38, closed symbols; X25-37, open-closed symbols; d, h: R35, closed symbols; R35-26, open symbols.

Methods: To determine survival following UV light exposure, 3×10^6 cells, suspended in 10 ml of phosphate-buffered-saline plus 10 mM glucose in an uncovered continuously shaken 60 mm Petri dish, were irradiated at room temperature from above with a germicidal lamp. Survival levels following chronic exposure to MM-C were determined by plating an appropriate number of cells in each concentration of MM-C. Colony forming ability was assessed at 10–14 days.

Fig. 3 Gel electrophoresis and blot hybridization to pSV2-GPT. **Methods:** 25 µg of cellular DNA was digested with restriction endonucleases overnight using 5 units of enzyme per µg of DNA. The DNA was then ethanol precipitated, resuspended in buffer (4% Ficoll, 0.012 M EDTA, 0.5% SDS and 0.01% bromophenol blue) and electrophoresed overnight on a horizontal 1% agarose gel. After treatment with alkaline and neutralization buffers, gels were transferred to nitrocellulose filters as described by Southern¹⁴ for 2.5 days. The filters were then baked for 2 h at 80 °C and pre-hybridized at 42 °C overnight in 20 ml of 50% formamide, 5×SSC, 5×Denhardt's solution, 0.02 M sodium phosphate (pH 6.5) and 200 µg ml⁻¹ sonicated salmon testes DNA (preheated to 100 °C for 5 min). Hybridization¹⁷ with ³²P-labelled pSV2-GPT (1×10⁶ c.p.m. ml⁻¹) was conducted in 10 ml of the same solution supplemented with 10% dextran sulphate overnight at 42 °C. High specific activity ³²P-DNA probe (2×10⁸ c.p.m. µg⁻¹ of DNA) was made by nick translation¹⁸. After hybridization, filters were washed twice at room temperature for 5 min and twice at 50 °C for 30 min in 2×SSC plus 0.1% SDS followed by two washes at 50 °C in 0.1×SSC plus 0.1% SDS. After air-drying, filters were exposed to Kodak XAR-5 film with one Cronex Lightning-Plus intensifying screen at -70 °C for periods up to 2 weeks.

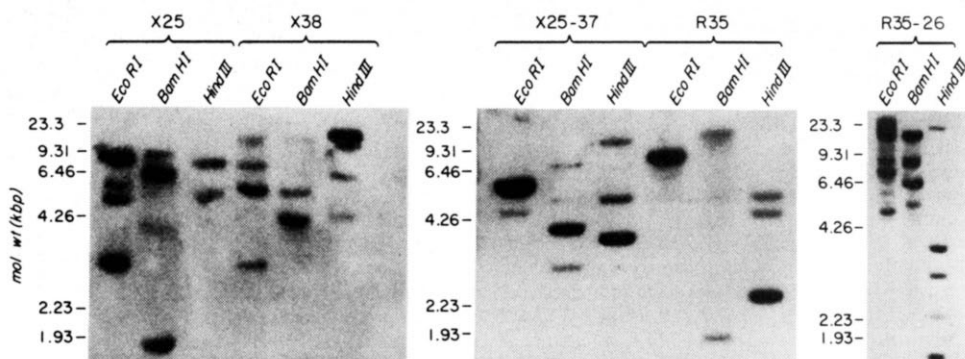
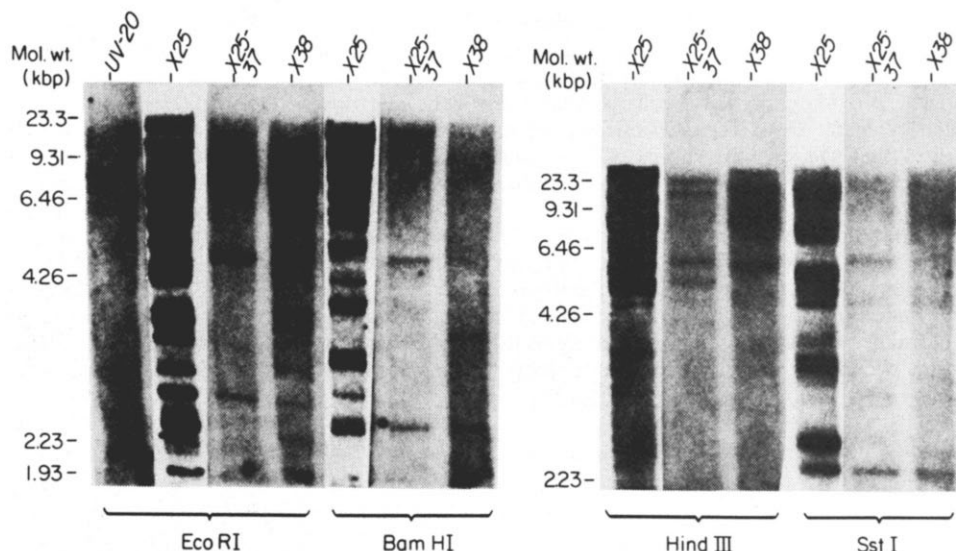


Fig. 4 Gel electrophoresis and blot hybridization to human cellular DNA. **Methods:** Conditions were the same as those described in the legend to Fig. 3 except that a vertical gel was used. ³²P-labelled human cellular DNA was used as a probe, and the final washes were conducted at 50–54 °C.



sequences as described in Fig. 4 by cleaving DNA with a variety of restriction endonucleases followed by gel electrophoresis and blot hybridization to a human cellular DNA probe.

As shown in Fig. 4, cellular DNA from transformant, X25, contains multiple restriction fragments that specifically hybridize to the human DNA probe, providing direct evidence that this transformant contains human DNA. However, these human sequences may only be present because selection for the expression of the *ecogpt* gene selects for a subpopulation of cells that are competent to uptake exogenous DNA. To determine whether any of these human sequences are associated with the DNA repair proficiency phenotype, cellular DNA from a secondary transformant, X25-37, was analysed in a similar manner. As shown in Fig. 4, this secondary transformant contains fewer human specific restriction fragments than the primary transformant due to the loss of sequences unrelated to the DNA repair proficient phenotype. More important, however, is the fact that both X25 and X25-37 contain a common set of restriction fragments. Figure 4 also demonstrates that the independently derived transformant, X38, shares some of these human specific sequences. These results indicate that independently derived transformants of the mutant cell line, UV-20, contain a common set of human specific restriction fragments, demonstrating the association of these sequences with a human DNA repair gene.

Although both 35 and R35-26, primary and secondary transformants derived from UV-41, contain pSV2-GPT sequences, results to date concerning the presence of human DNA have been inconclusive.

The results presented here provide both direct biological and molecular evidence for the DNA-mediated transfer of a human

DNA repair gene into repair deficient hamster mutants and a basis for the cloning and further characterization of the genes involved in human DNA repair systems. Although the relationship of this gene to any known human disorder is unknown, this work may provide new insights into the biochemical and genetic defects of a number of human DNA repair deficiency diseases.

We thank Dr L. Siminovich for comments on the manuscript and Mr Neil Amos and Mr George Cheong for technical assistance. This work was supported by the National Cancer Institute of Canada (NCI), the MRC of Canada and the Ontario Cancer Treatment and Research Foundation. J.S.R. and A.L.J. are recipients of Research Studentships from the NCI and the MRC, respectively.

Received 27 June; accepted 23 August 1983.

1. Friedberg, E. C., Ehman, V. K. & Williams, J. I. *Adv. radiat. Biol.* **8**, 85–174 (1979).
2. Thompson, L. H., Rubin, J. S., Cleaver, J. E., Whitmore, G. F. & Brookman, K. *Somat. Cell Genet.* **6**, 391–405 (1980).
3. Busch, D. B., Cleaver, J. E. & Glaser, D. A. *Somat. Cell Genet.* **6**, 407–418 (1980).
4. Thompson, L. H., Brookman, K. W., Dillenay, L. E., Mooney, C. L. & Carrano, A. V. *Somat. Cell Genet.* **8**, 759–773 (1982).
5. Thompson, L. H., Busch, D. B., Brookman, K., Mooney, C. L. & Glaser, D. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3734–3737 (1981).
6. Fujiwara, Y. & Tatsumi, M. *J. molec. Biol.* **113**, 635–649 (1977).
7. Gusella, J. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2829–2833 (1980).
8. Murray, M. J. *et al. Cell* **25**, 355–361 (1981).
9. Peruchio, M. *et al. Cell* **27**, 467–476 (1981).
10. Wigler, M. *et al. Cell* **16**, 777–785 (1979).
11. Mulligan, R. C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2072–2076 (1981).
12. Graham, F. L. & Van Der Eb, A. J. *Virology* **52**, 456–467 (1973).
13. Wigler, M., Pellicer, A., Silverstein, S. & Axel, R. *Cell* **14**, 725–731 (1978).
14. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
15. Takano, T., Noda, M. & Tamura, T. *Nature* **296**, 269–270 (1982).
16. Stanners, C. P., Eliceiri, G. L. & Green, H. *Nature* **230**, 52–54 (1971).
17. Wahl, G. M., Stein, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683–3687 (1979).
18. Rigby, P. W., Dieckman, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).

First mammal-like reptile from Australia

THE first mammal-like reptile fossil to be found in Australia, from the Arcadia Formation of south-east Queensland, was recently described by Thulborn¹. The specimen in question is a partial left quadrate whose morphology matches a suite of characters which Thulborn feels appear "to be unique to dicynodonts", following Watson's² description. Furthermore, Thulborn states that the specimen, Queensland Museum F12178, finds its closest match in the dicynodont *Kannemeyeria*, being virtually identical to one example illustrated by Watson². He therefore concludes that the quadrate belonged to a dicynodont closely related to (or perhaps identical with) *Kannemeyeria*.

This is an important conclusion because it is implicated in the dating of the Arcadia Formation which various evidence¹ points to being correlated with the *Lystrosaurus* Zone of South Africa. However, *Kannemeyeria* is usually thought to be associated with *Cynognathus* Zone (Lower Triassic) deposits, not with those of the earlier *Lystrosaurus* Zone. If the new fossil is indeed *Kannemeyeria* then either it is a very early occurrence of this genus, in the Arcadia Formation, or the Arcadia is not correlated with the South African *Lystrosaurus* Zone, despite the evidence to the contrary. Even if the fossil is only closely related to *Kannemeyeria*, not actually a member of that genus, this is still a surprising conclusion, as no known relatives of *Kannemeyeria* are found in deposits earlier than those of the *Cynognathus* Zone³.

However, there is little evidence to support the contention that the new specimen is *Kannemeyeria* or a similar genus, because the dicynodontian quadrate has no features of diagnostic significance (except perhaps in two advanced Triassic forms⁴). Although Queensland Museum F12178 is undoubtedly dicynodontian (*sensu* Cluver and King⁵), and although it and *Kannemeyeria* do appear similar, Queensland Museum F12178 also bears similarities to *Lystrosaurus*⁶, *Dicynodon*⁷, or indeed to almost any other dicynodontian. The problems of dating mentioned above may be solved by assuming that QM 12178 is not *Kannemeyeria*, or a close relative, but is in fact *Lystrosaurus*, the common fossil of the *Lystrosaurus* Zone.

Thulborn also mentions that *Kannemeyeria* and *Lystrosaurus* might have been contemporaneous in the Puesto Viejo Formation of Argentina, probably the earliest occurrence of *Kannemeyeria*. This is not a widely held view and has no fossil evidence to substantiate it, but it is in any case contradictory to Thulborn's suggestion that the Arcadia is not

necessarily correlated with the South African *Lystrosaurus* Zone. If *Kannemeyeria* and *Lystrosaurus* did coexist, there is nothing to prevent the Arcadia being correlated with the *Lystrosaurus* Zone. The Puesto Viejo cannot be used to illustrate this, however, as it is almost certainly of *Cynognathus* Zone age⁸.

In conclusion, it would seem simpler to regard the quadrate QM F12178 as belonging to a specimen of *Lystrosaurus*, and the Arcadia Formation as *Lystrosaurus* Zone age. This would be consistent with the evidence from associated remains and palynology mentioned by Thulborn.

G. M. KING

The Zoological Collections,
University Museum,
Parks Road,
Oxford OX1 3PW, UK

1. Thulborn, R. A. *Nature* **303**, 330–331 (1983).
2. Watson, D. M. S. *Proc. zool. Soc. Lond.* **118**, 823–877 (1948).
3. Anderson, J. M. & Cruickshank, A. R. I. *Palaeont. afr.* **21**, 15–44 (1978).
4. Keyser, A. W. *Palaeont. afr.* **17**, 57–68 (1974).
5. Cluver, M. A. & King, G. M. *Ann. S. Afr. Mus.* **91**, 195–273 (1983).
6. Cluver, M. A. *Ann. S. Afr. Mus.* **56**, 155–274 (1971).
7. King, G. M. *Phil. Trans. R. Soc. B291*, 243–322 (1981).
8. Bonaparte, J. F. *Proc. 2nd Gondwana Symp., South Africa*, 665–682 (C.S.I.R., Pretoria, 1970).

THULBORN REPLIES—King states that no known relatives of *Kannemeyeria* are found in deposits older than those of the *Cynognathus* Zone. However, Bonaparte reported¹ parts of several kannemeyeriid skulls, along with remains of a proterosuchid thecodontian, from the basal part of the Puesto Viejo Formation, Argentina. Radiometric dating confirmed that the fossiliferous beds were probably equivalent in age to the South African *Lystrosaurus* Zone.

To some extent dicynodont genera can be distinguished through differences in quadrate morphology—in the relative shapes and proportions of the condyles, in the degree of closure of the quadrate foramen and (in some cases) in the development of an antero-medial wing against the pterygoid. I compared the Arcadia specimen first-hand with the quadrates of various dicynodonts, including *Kannemeyeria* and African and Antarctic specimens of *Lystrosaurus*. In its size, and in its gross and detailed morphology, the Arcadia specimen found its closest match in the *Kannemeyeria* quadrate illustrated by Watson²; by contrast, it bears only a general resemblance to the quadrate of *Lystrosaurus* (see, for example, those specimens figured by Cluver³).

While the Arcadia quadrate is definitely identifiable as that of a dicynodont, it does not show sufficient detail to warrant its assignment to any new or existing genus,

be it *Kannemeyeria* or *Lystrosaurus*. It seems very likely that the Arcadia fauna is equivalent in age to that of the South African *Lystrosaurus* Zone^{4–6}, but this fact cannot be used to argue that any dicynodont fragment found in the Arcadia Formation must inevitably represent the genus *Lystrosaurus*.

1. Bonaparte, J. F. *Anais II Congr. Latino-Americano Paleont.*, Porto Alegre, Vol. 1, 277–288 (1981).
2. Watson, D. M. S. *Proc. zool. Soc. Lond.* **118**, 823–877 (1948).
3. Cluver, M. A. *Ann. S. Afr. Mus.* **56**, 155–274 (1971).
4. Foster, C. B. *Publ. geol. Surv. Qd* **372**, 1–244 (1979).
5. Thulborn, R. A. *Mem. Qd Mus.* **19**, 331–355 (1979).
6. Warren, A. *Alcheringa* **4**, 25–36 (1980).

RICHARD A. THULBORN

Department of Zoology,
University of Queensland,
St Lucia,
Queensland 4067, Australia

Correlation of Tulu Bor Tuff at Koobi Fora with the Sidi Hakoma Tuff at Hadar

BROWN¹ correlates the Sidi Hakoma Tuff (SHT) of Hadar with the Tulu Bor Tuff of East Turkana using chemical fingerprinting, and concludes that our K/Ar age² for the variably altered Kada Moumou Basalt (KMB), interpreted to be stratigraphically above SHT, is too old by ~0.5 Myr. This conclusion is hasty.

(1) Brown's statement that the mean of all type A KMB ages² is 3.12 Myr ignores our distinction between types A-1 (more altered) and A-2 (less altered), with mean ages of 2.94 ± 0.12 Myr and 3.33 ± 0.06 Myr, respectively. This evidently resulted from his not realizing that KM-1-74 (lab) and KM-2-74 (lab) were re-labelled from KM-2-74 and KM-1-74 of ref. 3, when we discovered that the latter had been accidentally interchanged after grinding⁴. Brown's correlation places KMB in the Mammoth Event (3.15–3.05 Myr); if this is true, then not only is the unique type B too old, but so is the common A-2 variety. Our oversight in not including the older age measurements obtained for the type A-1 sample in the density fractionation experiment² does blur the age distinction between types A-1 and A-2 somewhat (F. H. Brown, personal communication).

By this reasoning it becomes fortuitous that the most altered petrographic type yields the best K/Ar age, while the progressively less altered types give ages too old because of extraneous ⁴⁰Ar*. Just considering the internal systematics of the KMB data and the lack of geological evidence for extraneous ⁴⁰Ar* leads to the conclusion that the

3.61 ± 0.13 Myr result on the least altered material is a reasonable minimum age. If the KMB-B age turns out to be too high then the K/Ar data indicate that this flow is unusual, having both extraneous $^{40}\text{Ar}^*$ and alteration.

(2) Volcanic glass from separate eruptions of large caldera collapse centres can have very similar chemistries, although separated in eruption by several hundred thousand years. For example, the 2.0-Myr and 0.6-Myr eruptions of Yellowstone (USA) match the similarity of SHT and Tulu Bor using more elements than Brown did⁵⁻⁸. Ultimately only primary geochronology could distinguish between these tuffs⁹. Brown provides strong evidence that SHT and Tulu Bor are from the same magma sources, but whether they are also from the same eruption remains inconclusive.

Considering the rift basin setting of SHT and Tulu Bor, 900 km apart, ash from this large magnitude eruption should be preserved somewhere between the two sites. Establishing correlations in the inter-basin area will confirm the SHT-Tulu Bor correlation whereas if two stratigraphically distinct, but chemically similar, tuffs are found, it will be necessary to reject it.

(3) An unusual criterion of SHT should be sought at Turkana. Stratigraphically superjacent to SHT is a volcanoclastic sand, rich in sanidine (mode = 15%) compared with uncommon K-feldspar in typical Hadar sands (mode = 0.6–1.2%). The bulk of this sand's sanidine produces concordant ages at 9.00 ± 0.14 Myr (new constants) despite a range in K_2O from 3.9 to 7.4% (ref. 10). Did the SHT eruption disperse an old pre-existing roof from the volcanic centre?

(4) If SHT and Tulu Bor do correlate another possibility is that our older age structure for Hadar, based on KMB-B and BKT-2, is more appropriate than that proposed by Brown for Turkana, which is heavily weighted by magnetic stratigraphy. Most K/Ar data for the lower Omo and East Turkana sections have been imprecise, except for Shungura B-10, the K/Ar date of which Brown rejects as too old. Palaeomagnetic stratigraphy is part of the overall criteria to establish age but, alone, is prone to errors in polarity and in recognizing whether disconformities have lost reversals. As with the KBS tuff, further geochronology should resolve which age framework is right, and whether SHT and Tulu Bor correlate.

We are pleased by Brown's discovery that the same volcanic centre probably shed ash, perhaps in the same eruption, to both the Awash and Turkana Basins. Faunal records from both basins are largely complementary. Establishing the same tephra time-plane across both basins would represent a fortunate circumstance for understanding East African faunal evolution. We need to locate the source volcano that could have produced a voluminous SHT-Tulu Bor ash, and

unravel its eruptive history. These hypotheses are testable and point out the difficulties in accepting *a priori* Brown's correlation.

We thank G. A. Izett, A. Sarna-Wojcicki and J. A. Westgate for unpublished data on North American caldera eruptions.

J. L. ARONSON

*Department of Geological Sciences,
Case Western Reserve University,
Cleveland, Ohio 44106, USA*

R. C. WALTER

*Department of Geology and
Geophysics,
University of Toronto,
Toronto, Ontario, Canada M5S 1A1.*

M. TAIEB

*Laboratoire de Géologie
du Quaternaire,
Centre Nationale de la
Recherche Scientifique,
132 88 Marseille, Cedex 2, France*

1. Brown, F. *Nature* **300**, 631–633 (1982).
2. Walter, R. C. & Aronson, J. L. *Nature* **296**, 122–127 (1982).
3. Aronson, J. L. *et al. Nature* **267**, 323–327 (1977).
4. Walter, R. C. thesis, Case Western Reserve Univ. (1980).
5. Izett, G. A. *J. geophys. Res.* **86**, 10200–10222 (1981).
6. Izett, G. A. *US geol. Surv. Open File Rep.* 82-582, 44 (1982).
7. Sarna-Wojcicki, A. M. *et al. US geol. Surv. Open File Rep.* 80-231 (1980).
8. Westgate, J. A., Christiansen, E. A. & Boellstorff, J. D. *Can. J. Earth Sci.* **14**, 357–374 (1977).
9. Naeser, C. W., Izett, G. A. & Wilcox, R. E. *Geology* **1**, 187–189 (1973).
10. Aronson, J. L., Walter, R. C., Taieb, M. & Naeser, C. W. *Proc. 8th Panafri. Congr. of Prehistory and Quat. Studies, Nairobi, 1977*, 47–52 (1980).

BROWN REPLIES—There are two separate issues addressed above—correlation and chronology. I agree that geochemical correlation is permissive, but, as noted earlier, the magnetic polarity record and faunal data^{1,2} also support the correlation. I have recently determined that the SHT contains about 5% of shards which correspond to the α -Tulu Bor composition. This is also characteristic of the β -Tulu Bor³, hence I believe the correlation is secure. Note that the Tulu Bor Tuff (70 samples) is compositionally quite distinct from 110 other tuffs in the Turkana Basin (475 samples), and is unusual for East African silicic volcanics in general.

The distribution of the Tulu Bor and/or SHT tuffs does not necessarily imply an excessively large eruption. In both cases the tuffs are water-deposited from rivers^{4,5}. It may only be necessary that ash was supplied to the upper reaches of the Awash and Omo rivers, which have a common drainage divide.

Concerning the chronology, both parties agree that the SHT and Tulu Bor Tuffs are between 3.2 and 4 Myr old, and of normal polarity⁶. There are thus two possible placements—3.17–3.41 Myr (bottom of Gauss Chron) or 3.82–3.92 Myr

(Cochiti Subchron). The magnetostratigraphy in section A⁷ at Hadar, and in the Turkana Basin, favours the younger placement. The older placement requires slow sedimentation rates in the stratigraphic interval between the Tulu Bor and the next highest dated horizon, and between the SHT and the bottom of the reversed interval identified as the Mammoth Subchron in Section A⁷. Alternatively, disconformities might exist in both sections in this interval, but none has been documented. Still, the earlier placement cannot be categorically denied given the poor radiometric control on the earlier part of the Turkana sequence at present.

The older placement is favoured by the age on the type B KMB of 3.61 Myr, only provided that age is meaningful and the KMB is placed correctly in the section. I admit missing the footnote⁶ which equates KM-1-74(lab) with KM-2-74. Acceptance of 3.61 Myr as a minimum age on KMB depends strongly on petrographic arguments, thus I find it peculiar that the mix-up was not dealt with in the revised data⁸, especially as KM-1-74 had been discussed previously⁷ as type A-1 and KM-2-74 as type A-2. Even so, type A-1 KMB sometimes yields ages older than expected, and type B KMB sometimes yields ages younger than expected⁸. The reason for this is unclear; the reported inverse correlation of age with glass content in type A-1 (ref. 8) is very weak.

The Hadar faunas are more securely related to the SHT than to the KMB (Fig. 2 of ref. 5). Placement of the KMB in the Hadar section has been difficult. It seems possible that the KMB caps an older sedimentary sequence separated from the main Hadar sequence by an unconformity. This possibility might explain the difficulty in tracing beds below the basalt into the main Hadar sequence (see Fig. 4 of ref. 9 and Figs 34:3 and 34:4 of ref. 5). Aronson, Walter, Taieb and Tiercelin considered the possibility of the Kadada Moumou Plateau as a horst, but rejected this in favour of a minor intraformational unconformity above KMB (J. Aronson, personal communication).

FRANCIS H. BROWN

*Department of Geology and Geophysics,
University of Utah,
Salt Lake City, Utah 84112, USA*

1. Boaz, N. T., Howell, F. C. & McCrossin, M. L. *Nature* **300**, 633–635 (1982).
2. Cooke, H. B. S. *Kirlandia* **29**, 1–63 (1978).
3. Cerling, T. E. & Brown, F. H. *Nature* **299**, 216–221 (1982).
4. Brown, F. H. *Calibration of Hominoid Evolution* (eds Bishop, W. W. & Miller, J. A.) 273–287 (Scottish Academic, Edinburgh, 1972).
5. Johanson, D. C., Taieb, M., Gray, B. T. & Coppens, Y. *Geological Background to Fossil Man* (ed. Bishop, W. W.) 549–564 (Scottish Academic, Edinburgh, 1978).
6. Walter, R. C. thesis, Case Western Reserve Univ. (1980).
7. Aronson, J. L. *et al. Nature* **267**, 323–327 (1977).
8. Walter, R. C. & Aronson, J. L. *Nature* **296**, 122–127 (1982).
9. Taieb, M., Johanson, D. C., Coppens, Y. & Aronson, J. L. *Nature* **260**, 289–293 (1976).